(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 February 2003 (27.02.2003)

PCT

(10) International Publication Number
WO 03/016535 A2

(51) International Patent Classification[7]: C12N 15/56, 9/30, C07K 14/38, 16/14

(21) International Application Number: PCT/NL02/00522

(22) International Filing Date: 2 August 2002 (02.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

01000379.6	16 August 2001 (16.08.2001)	EP
01000380.4	16 August 2001 (16.08.2001)	EP
01000381.2	16 August 2001 (16.08.2001)	EP
01000382.0	16 August 2001 (16.08.2001)	EP
01000383.8	16 August 2001 (16.08.2001)	EP
01000384.6	16 August 2001 (16.08.2001)	EP

(71) Applicant *(for all designated States except US)*: **DSM N.V.** [NL/NL]; Het Overloon, NL-6411 TE Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MAIER, Dieter** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE). **STOCK, Alexander** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE). **WAGNER, Christian** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE). **FOLKERS, Ulrike** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE). **ALBERMANN, Kaj** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE). **HOPPER, Sylvia** [DE/DE]; Lochhamer Str. 11, 82152 Martinsried (DE).

(74) Agent: **HABETS, Winand,Johannes, Antonius**; DSM Patents & Trademarks, P.O. Box 9, NL-6160 MA Geleen (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/016535 A2

(54) Title: NOVEL AMYLASES AND USES THEREOF

(57) Abstract: The invention relates to newly identified polynucleotide sequences comprising genes that encode novel amylases isolated from Aspergillus niger. The invention features the full length nucleotide sequences of the novel genes, the cDNA sequences comprising the full length coding sequence of the novel amylases as well as the amino acid sequence of the full-length functional proteins and functional equivalents thereof. The invention also relates to methods of using these enzymes in industrial processes and methods of diagnosing fungal infections. Also included in the invention are cells transformed with a polynucleotide according to the invention and cells wherein an amylase according to the invention is genetically modified to enhance or reduce its activity and/or level of expression.

NOVEL AMYLASES AND USES THEREOF

Field of the invention

The invention relates to newly identified polynucleotide sequences comprising genes that encode novel amylases isolated from Aspergillus niger. The invention features the full length nucleotide sequence of the novel genes, the cDNA sequence comprising the full length coding sequences of the novel amylases as well as the amino acid sequences of the full-length functional proteins and functional equivalents thereof. The invention also relates to methods of using these enzymes in industrial processes and methods of diagnosing fungal infections. Also included in the invention are cells transformed with a polynucleotide according to the invention and cells wherein an amylase according to the invention is genetically modified to enhance its activity and/or level of expression.

Background of the invention

Industrial processes for the hydrolysis of starch to glucose rely on inorganic acids or enzyme catalysis. The use of enzymes is preferred currently and offers a number of advantages associated with improved yields and favourable economics. Enzymatic hydrolysis allows greater control over amylolysis, the specifity of the reaction, and the stability of the generated products. The milder reaction conditions involve lower temperatures and near-neutral pH, thus reducing unwanted side reactions. Fewer off-flavor and off-color compounds are produced, especially 5-hydroxy-2-methylfurfuraldehyde, anhydroglucose compounds, and undesirable salts. Enzymatic methods are favored because they also lower energy requirements and eliminate neutralization steps.

Alpha- amylases (E.C. 3.2.1.1) or α-amylases catalyse the endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. They are also known as 1,4-alpha-D-glucan glucanohydrolase, Taka-amylase, endoamylase or glycogenase. Alpha amylases act on starch, glycogen and related polysaccharides and oligosaccharides in a random manner; reducing groups are liberated in the alpha-configuration.

Beta-amylases (E.C.3.2.1.2) catalyse the hydrolysis of 1,4-alpha-glucosidic linkages in polysaccharides so as to remove successive maltose units from

the non-reducing ends of the chains. Other names are: 1,4-alpha-D-glucan maltohydrolase, Saccharogen amylase, Glycogenase. Beta amylases act on starch, glycogen and related polysaccharides and oligosaccharides producing beta-maltose by an inversion.

Glucoamylases (E.C.3.2.1.3) catalyse the hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose. *Other names are: Glucan 1,4-alpha-glucosidase.* 1,4-alpha-D-glucan glucohydrolase. Amyloglucosidase. Gamma-amylase. Lysosomal alpha-glucosidase. Exo-1,4-alpha-glucosidase. Most forms of the enzyme can rapidly hydrolyse 1,6-alpha-D-glucosidic bonds when the next bond in sequence is 1,4, and some preparations of this enzyme hydrolyse 1,6- and 1,3-alpha-D-glucosidic bonds in other polysaccharides.

Amylases may conviently be produced in microorganisms. Microbial amylases are available from a variety of sources; Bacillus spec. are a common source of bacterial enzymes, whereas fungal enzymes are commonly produced in Aspergillus spec.

The low pH optimum of most fungal amylases permits the convenient use of acid conditions for the saccharification. Such conditions reduce unwanted isomerization reactions to fructose and other sugars that may reduce the glucose yield. Moreover, acid conditions restrict the growth of contaminating microorganisms in the saccharification reactors.

Amylases may be used in a manifold of industrial applications, including baking, brewing, the production of corn syrup and alcohol as well as in vinegar fermentation.

Malted wheat, barley, bacteria, and fungi are typical sources of α-amylase for baking purposes. Fungal α-amylase is added to bread doughs in the form of diluted powders, prepacked doses, or water dispersible tablets. The enzyme may be added to flours at the bakery or, more rarely, at the mill itself. Malted wheat and barley also can serve as sources of amylolytic activity when flours from these grains are blended with the final product at the mill. The properties of bacterial α-amylase permit its application to the production of coffee cake, fruit cake, brownies, cookies, snacks, and crackers. Fungal α-amylase, usually from *A. oryzae, A. niger, A. awamori,* or species of *Rhizopus*, is used to supplement the amylolytic activity in flour. Enzymes from these sources can raise the levels of fermentable monosaccharides and disaccharides of dough from a native level of 0.5% to concentrations that promote

yeast growth. The sustained release of glucose and maltose by added fungal and endogenous enzymes provides the nutrients essential for yeast metabolism and gas production during panary fermentation. The *A. oryzae* α-amylase is sometimes favored for baking applications over the bacterial enzyme obtained from *Bacillus* species since the fungal enzyme is heat labile at 60-70 °C and does not survive the baking process. Its thermolability prevents enzymatic action on the gelatinised starch in the finished loaf which would cause a soft or sticky crumb. Bacterial α-amylase is also used with good results, but its dose must be measured carefully to avoid a bread with a gummy mouthfeel. Amylase supplementation is also beneficial and sometimes essential, since white bread flours contain 6.7-10.5% damaged starch. The added enzyme degrades damaged, ruptured starch granules that usually are present in bread flour more efficiently than does wheat β-amylase (Bigelis R. in: Enzymes in Food processing, Nagodawithana and Reed Eds. Acad. Press Inc p121-158 and references cited therein).

Amylase supplementation can improve other characteristics of bread quality, in addition to improving the quality of rolls, buns, and crackers, when used during manufacturing processes for these baked goods. In bread baking, treatment with fungal or bacterial amylase lowers the viscosity of bread dough, thereby improving the ease of manipulation by manual workers or machines. Measured doses of enzyme also lower the compressibility of the loaf, producing a softer bread. Further, such processing increases the bread volume by reducing the viscosity of the gelling starch and allowing greater expansion during baking before protein denaturation and enzyme inactivation fix the volume of the loaf. Favorable effects on taste, crust properties, and toasting qualities are observed. The storage characteristics of breads are changed also, yielding a product with a softer, more compressible crumb that firms more slowly and keeps longer, as determined by taste panels. Amylolytic activity also may elevate the sugar concentration in bread and yield a preferred sweeter product with sensory advantages (Bigelis R. in: Enzymes in Food processing, Nagodawithana and Reed Eds. Acad. Press Inc p121-158 and references cited therein).

In brewing, added enzymes contribute to the action of endogenous barley β-amylase and aid in the starch digestion process. Such added enzymes are especially important when nonmalted cereal grains such as corn and rice, termed adjuncts, are used. Since these adjunct grains are deficient in carbohydrases, fungal α-amylase and glucoamylase can increase starch digestion, reduce the proportion of unmalted grain, and insure a consistent quality of the mash. Amylase solubilizes barley

amylose and amylopectin, exposing these substrates to further degradation by barley β-amylase. As a result, the levels of maltose and small dextrins are raised, eventually yielding the wort ingredients that promote yeast fermentation. Amylase preparations with low transglucosidase activity are favored since trace levels of this enzyme generate isomaltose and panose, both of which are nonfermentable by yeast. The source of amylase activity for brewing applications is generally enzyme from *Aspergillus* species such as *A. niger* or *A. oryzae*. Protease from these sources may be added in concert with amylase to solubilize protein and release amino acids essential for yeast proliferation . (Bigelis R. in: Enzymes in Food processing, Nagodawithana and Reed Eds. Acad. Press Inc p121-158 and references cited therein)

In the above processes, it is advantageous to use enzymes that are obtained by recombinant DNA techniques. Such recombinant enzymes have a number of advantages over their traditionally purified counterparts. Recombinant enzymes may be produced at a low cost price, high yield, free from contaminating agents like bacteria or viruses but also free from bacterial toxins or contaminating other enzyme activities.

Molecular cloning of amylases in fungi has been described. The DNA and deduced amino acid sequences of certain alpha-amylases from Aspergillus oryzae, A. niger and A. shirousamii are given in Wirsel et al. Mol. Microbiol 1989, (1) 3-14, Boel et al., Biochemistry 1990 (29) 6244 –6249, and Shibuya et al., Biosci. Biotech. Biochem. 1992 (56) 174-179. Molecular cloning of an α-amylase from Bacillus amyloliquefaciens is described by Takkinen et al. J. Biol. Chem. 1983, (258) 1007-1013.

It is important that amylases, in particular α-amylases, can be produced at low costs. This may be achieved by improving the production efficiency (higher expression levels) or by providing enzymes with an improved specific activity (higher activity per mg of enzyme). It is therefore an object of the present invention to provide improved enzymes with an improved production efficiency and/or improved specific activity.

When α amylases are used as bread improvers, it is advantageous to provide them, preferably together with other enzymes, in a liquid preparation. Enzyme stabilisers like glycerol are a major cost factor of liquid bread improvers and consequently there is a need for more stable α-amylases for use in such preparations in order to lower the amount of stabilisers. It is also an object of the present invention to provide more stable α-amylases.

α-Amylases are often used in combination with ascorbic acid, which

tends to become unstable at higher pH values. α-Amylases on the other hand become unstable at lower pH values. As a compromise between the two requirements, such preparations are usually kept at a pH value around 4.7. It would therefore be advantageous to have α-amylases with a lower pH optimum and/or a higher stability at low pH values, preferably below pH 4.7. The present invention provides such enzymes.

Ascorbic acid is used in combination with α-amylases in many applications where it is converted into a number of chelating agents, e.g. oxalate. Oxalate is able to bind Ca ions and since α-amylases require Ca ions for their stability, oxalate acts as a destabiliser for these α-amylase enzyme preparations. It is therefore an object underlying the present invention to provide enzymes that are less dependent on Ca ions for their stability.

Another characteristic of α-amylases according to the prior art is their limited thermostability. Fungal α-amylases are inactivated at about 65 °C, therefore they are heat-inactivated at the beginning of the baking process. Also, this property makes fungal α-amylases unsuited for activity measurements in the Hagberg falling number method (AACC, 1983, Method 56-81 A) and the Brabender amylograph method (AACC, 1983, Method 22-1). Also, prolonged storage at temperatures slightly above room temperature sometimes deteriorates enzyme activity. It is therefore an object of the present invention to provide α-amylases with improved thermostability.

Object of the invention

It is an object of the invention to provide novel polynucleotides encoding improved novel amylases. A further object is to provide improved naturally and recombinantly produced amylases as well as recombinant strains producing these. Also fusion polypeptides are part of the invention as well as methods of making and using the polynucleotides and polypeptides according to the invention.

Summary of the invention

The invention relates to isolated polypeptides having α-amylase activity and one or more characteristics selected from the group consisting of:

1) An isolated polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof,
2) An isolated polypeptide obtainable by expressing a polynucleotide having a

nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or a vector comprising said polynucleotides or functional equivalents thereof in an appropriate host cell, e.g. Aspergillus niger.

3) Polypeptide comprising a functional domain of a polypeptide according to (1) or (2)
4) An allelic variant of (1), (2) or (3),
5) A fragment of (1), (2), (3) or (4)
6) A polypeptide having improved specific activity and/or improved production efficiency expressed as enzyme activity per mg of purified enzyme or as enzyme activity per ml culture volume or per mg of biomass produced
7) A polypeptide with improved stability, preferably stable in the presence of less than 50% glycerol, preferably less than 40% glycerol, more preferably less than 30% glycerol, more preferably less than 20% glycerol, more preferably less than 10% glycerol, most preferably in the absence of glycerol
8) A polypeptide stable at pH values below 4.7, preferably below pH 4.0, even more preferably below pH 3.5,
9) A polypeptide with improved stability towards Ca ions.

It is expressly mentioned that α-amylases according to the invention may have one or more of the above characteristics. Methods for determining specific activity, production efficiency, stability, pH optimum and acid stability are well known in the art. Among others they may be found in the materials and methods section of WO 00/60058.

The invention also relates to polynucleotides encoding any of the polypeptides mentioned above.

More in particular, the invention provides for polynucleotides having a nucleotide sequence that hybridises preferably under highly stringent conditions to a sequence having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12. Consequently, the invention provides nucleic acids that are about 40%, preferably 65%, more preferably 70%, even more

preferably 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homologous to any sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12.

In a more preferred embodiment the invention provides for such an isolated polynucleotide obtainable from a filamentous fungus, in particular A. niger is preferred.

In one embodiment, the invention provides for an isolated polynucleotide comprising a nucleic acid sequence encoding a polypeptide with having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

In a further preferred embodiment, the invention provides an isolated polynucleotide encoding at least one functional domain of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

In a preferred embodiment the invention provides an amylase gene having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6. In another aspect the invention provides a polynucleotide, preferably a cDNA encoding an A. niger amylase having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or variants or fragments of that polypeptide. In a preferred embodiment the cDNA has a nucleotide sequence selected from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or functional equivalents thereof.

In an even further preferred embodiment, the invention provides for a polynucleotide comprising the coding sequence of the polynucleotides according to the invention, preferred is a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12.

The invention also relates to vectors comprising a polynucleotide

sequence according to the invention and primers, probes and fragments that may be used to amplify or detect the DNA according to the invention.

In a further preferred embodiment, a vector is provided wherein the polynucleotide sequence according to the invention is functionally linked with regulatory sequences suitable for expression of the encoded amino acid sequence in a suitable host cell, such as A. niger or A. oryzea. The invention also provides methods for preparing polynucleotides and vectors according to the invention.

The invention also relates to recombinantly produced host cells that contain heterologous or homologous polynucleotides according to the invention.

In another embodiment, the invention provides recombinant host cells wherein the expression of an amylase according to the invention is significantly increased or wherein the activity of the amylase is increased.

In another embodiment the invention provides for a recombinantly produced host cell that contains heterologous or homologous DNA according to the invention and wherein the cell is capable of producing a functional amylase according to the invention, preferably a cell capable of over-expressing the amylase according to the invention, for example an Aspergillus strain comprising an increased copy number of a gene or cDNA according to the invention.

In yet another aspect of the invention, a purified polypeptide is provided. The polypeptides according to the invention include the polypeptides encoded by the polynucleotides according to the invention. Especially preferred is a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

Fusion proteins comprising a polypeptide according to the invention are also within the scope of the invention. The invention also provides methods of making the polypeptides according to the invention.

The invention also relates to the use of the amylase according to the invention in any industrial process as described herein

Detailed description of the invention

Polynucleotides

The present invention provides polynucleotides encoding an alpha-amylase having an amino acid sequence selected from the group consisting of SEQ ID

NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof. The sequence of the genes encoding a protein according to the invention was determined by sequencing a genomic clone obtained from Aspergillus niger. The invention provides polynucleotide sequences comprising the gene encoding the A niger alpha amylase as well as its complete cDNA sequence and its coding sequence. Accordingly, the invention relates to an isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or functional equivalents thereof.

More in particular, the invention relates to an isolated polynucleotide hybridisable under stringent conditions, preferably under highly stringent conditions, to a polynucleotide having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12. Advantageously, such polynucleotides may be obtained from filamentous fungi, in particular from Aspergillus niger. More specifically, the invention relates to an isolated polynucleotide having a nucleotide sequence having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12.

The invention also relates to an isolated polynucleotide encoding at least one functional domain of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which may be isolated from chromosomal DNA, which include an open reading frame encoding a protein, e.g. an A. niger amylase. A gene may include coding sequences, non-coding sequences, introns and regulatory sequences. Moreover, a gene refers to an isolated nucleic acid molecule as defined herein.

A nucleic acid molecule of the present invention, such as a nucleic acid molecule having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID

NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or a functional equivalent thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or portion of the nucleic acid sequence of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 as a hybridization probe, nucleic acid molecules according to the invention can be isolated using standard hybridization and cloning techniques (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence information provided herein.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

Furthermore, oligonucleotides corresponding to or hybridisable to nucleotide sequences according to the invention can be prepared by standard synthetic techniques, e. g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12. The sequence information provided in SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 corresponds to the coding region of the A. niger alpha amylases genes provided in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 respectively. This cDNA comprises sequences encoding the A. niger alpha amylases having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 respectively.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the

nucleotide sequence shown in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or a functional equivalent of these nucleotide sequences.

A nucleic acid molecule which is complementary to another nucleotide sequence is one which is sufficiently complementary to the other nucleotide sequence such that it can hybridize to the other nucleotide sequence thereby forming a stable duplex.

One aspect of the invention pertains to isolated nucleic acid molecules that encode a polypeptide of the invention or a functional equivalent thereof such as a biologically active fragment or domain, as well as nucleic acid molecules sufficient for use as hybridisation probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules.

An "isolated polynucleotide" or "isolated nucleic acid" is a DNA or RNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. Thus, in one embodiment, an isolated nucleic acid includes some or all of the 5' non-coding (e.g., promotor) sequences that are immediately contiguous to the coding sequence. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide that is substantially free of cellular material, viral material, or culture medium (when produced by recombinant DNA techniques), or chemical precursors or other chemicals (when chemically synthesized). Moreover, an "isolated nucleic acid fragment" is a nucleic acid fragment that is not naturally occurring as a fragment and would not be found in the natural state.

As used herein, the terms "polynucleotide" or "nucleic acid molecule" are intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. The nucleic acid may be synthesized using

oligonucleotide analogs or derivatives (e.g., inosine or phosphorothioate nucleotides). Such oligonucleotides can be used, for example, to prepare nucleic acids that have altered base-pairing abilities or increased resistance to nucleases.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a nucleic acid molecule according to the invention. Also included within the scope of the invention are the complement strands of the nucleic acid molecules described herein.

In certain applications it may be advantageous to have products that are free of amylase activity. Such products may be produced in microorganisms wherein an amylase gene according to the invention is eliminated or wherein its activity is reduced. Such microorganisms may be obtained by recombinant DNA technology, for instance by knocking out the expression of a gene according to the invention. Amylase deficient mutants may be advantageously used for the production of milk clotting enzymes where contamination with amylases is undesired.

Instead of elimination of amylase activities via disruption or mutagenesis, reduced amylase activity can also be achieved via down-regulation of the amylase activities. This may be achieved by genetically altering the promoter or other regulatory sequences of the gene(s) according to the invention. With the help of the sequence information provided herein, the skilled person will know how to achieve the goal of providing mutant microorganisms with reduced or eliminated amylase activity.

Sequencing errors

The sequence information as provided herein should not be so narrowly construed as to require inclusion of erroneously identified bases. The specific sequences disclosed herein can be readily used to isolate the complete gene from filamentous fungi, in particular A. niger which in turn can easily be subjected to further sequence analyses thereby identifying sequencing errors.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the

actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

The person skilled in the art is capable of identifying such erroneously identified bases and knows how to correct for such errors.

Nucleic acid fragments, probes and primers

A nucleic acid molecule according to the invention may comprise only a portion or a fragment of the nucleic acid sequence shown in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12, for example a fragment which can be used as a probe or primer or a fragment encoding a portion of a protein according to the invention. The nucleotide sequence determined from the cloning of the alpha amylase gene and cDNA allows for the generation of probes and primers designed for use in identifying and/or cloning other alpha amylase family members, as well as homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide which typically comprises a region of nucleotide sequence that hybridizes preferably under highly stringent conditions to at least about 12 or 15, preferably about 18 or 20, preferably about 22 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 or more consecutive nucleotides of a nucleotide sequence shown in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or of a functional equivalent thereof.

Probes based on the nucleotide sequences provided herein can be used to detect transcripts or genomic sequences encoding the same or homologous proteins for instance in other organisms. In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme cofactor. Such probes can also be used as part of a diagnostic test kit for identifying cells which express an alpha-

amylase.

Identity & homology

The terms "homology" or "percent identity" are used interchangeably herein. For the purpose of this invention, it is defined here that in order to determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino acid sequence or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions/total number of positions (i.e. overlapping positions) x 100). Preferably, the two sequences are the same length.

The skilled person will be aware of the fact that several different computer programms are available to determine the homology between two sequences. For instance, a comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (formerly available at http://www.gcg.com now at http://www.accelrys.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (formerly available at http://www.gcg.com now at http://www.accelrys.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity two amino acid or nucleotide sequence is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)

which has been incorporated into the ALIGN program (version 2.0) (available at: http://vega.igh.cnrs.fr/bin/align-guess.cgi) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J. Mol. Biol. 215:403—10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

Hybridisation

As used herein, the term "hybridizing" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about 50%, at least about 40%, at least about 70%, more preferably at least about 80%, even more preferably at least about 85% to 90%, more preferably at least 95% homologous to each other typically remain hybridized to each other.

A preferred, non-limiting example of such hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45 °C, followed by one or more washes in 1 X SSC, 0.1 % SDS at 50 °C, preferably at 55 °C, preferably at 60 °C and even more preferably at 65 °C.

Highly stringent conditions include, for example, hybridizing at 68 °C in 5x SSC/5x Denhardt's solution / 1.0% SDS and washing in 0.2x SSC/0.1% SDS at room temperature. Alternatively, washing may be performed at 42 °C.

The skilled artisan will know which conditions to apply for stringent and highly stringent hybridisation conditions. Additional guidance regarding such conditions is readily available in the art, for example, in Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.).

Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of mRNAs), or to a complementary stretch of T (or U) resides, would not be included in a polynucleotide of the invention used to specifically hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-standed cDNA clone).

Obtaining full length DNA from other organisms

In a typical approach, cDNA libraries constructed from other organisms, e.g. filamentous fungi, in particular from the species Aspergillus can be screened.

For example, Aspergillus strains can be screened for homologous polynucleotides by Northern blot analysis. Upon detection of transcripts homologous to polynucleotides according to the invention, cDNA libraries can be constructed from RNA isolated from the appropriate strain, utilizing standard techniques well known to those of skill in the art. Alternatively, a total genomic DNA library can be screened using a probe hybridisable to a polynucleotide according to the invention.

Homologous gene sequences can be isolated, for example, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of nucleotide sequences as taught herein.

The template for the reaction can be cDNA obtained by reverse transcription of mRNA prepared from strains known or suspected to express a polynucleotide according to the invention. The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a new alpha amylase nucleic acid sequence, or a functional equivalent thereof.

The PCR fragment can then be used to isolate a full length cDNA clone by a variety of known methods. For example, the amplified fragment can be labeled and used to screen a bacteriophage or cosmid cDNA library. Alternatively, the labeled fragment can be used to screen a genomic library.

PCR technology also can be used to isolate full length cDNA sequences from other organisms. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis.

The resulting RNA/DNA hybrid can then be "tailed" (e.g., with guanines) using a standard terminal transferase reaction, the hybrid can be digested with RNase H, and second strand synthesis can then be primed (e.g., with a poly-C primer). Thus, cDNA sequences upstream of the amplified fragment can easily be isolated. For a review of useful cloning strategies, see e.g.,Sambrook et al., supra; and Ausubel et al., supra.

Vectors

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a protein according to the invention or a functional equivalent thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. The terms "plasmid" and "vector" can be used interchangeably herein as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vector includes one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operatively linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner

which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signal). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in a certain host cell (e.g. tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, encoded by nucleic acids as described herein (e.g. alpha-amylases, mutant alpha amylases, fragments thereof, variants or functional equivalents thereof, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of alpha amylases in prokaryotic or eukaryotic cells. For example, a protein according to the invention can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episome, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled person. In a specific embodiment, promoters are preferred that are capable of directing a high expression

level of amylases in filamentous fungi. Such promoters are known in the art. The expression constructs may contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-percipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipidmediated transfection or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual, 2nd,ed. Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., *Basic Methods in Molecular Biology* (1986) and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methatrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a protein according to the invention or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g. cells that have incorporated the selectable marker gene will survive, while the other cells die).

Expression of proteins in prokaryotes is often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, e.g. to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification.

Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognation sequences, include Factor Xa, thrombin and enterokinase.

As indicated, the expression vectors will preferably contain selectable markers. Such markers include dihydrofolate reductase or neomycin resistance for eukarotic cell culture and tetracyline or ampicilling resistance for culturing in *E. coli* and other bacteria. Representative examples of appropriate host include bacterial cells, such as *E. coli*, Streptomyces and Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS and Bowes melanoma; and plant cells. Appropriate culture media and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria are pQE70, pQE60 and PQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16A, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are PWLNEO, pSV2CAT, pOG44, pZT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Among known bacterial promotors for use in the present invention include *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR, PL promoters and the trp promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

For secretion of the translated protein into the lumen of the

endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretation signal may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification.

Polypeptides according to the invention

The invention provides an isolated polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18, an amino acid sequence obtainable by expressing the polynucleotide of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 in an appropriate host. Also, a peptide or polypeptide comprising a functional equivalent of the above polypeptides is comprised within the present invention. The above polypeptides are collectively comprised in the term "polypeptides according to the invention"

The terms "peptide" and "oligopeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of at least two amino acids coupled by peptidyl linkages. The word "polypeptide" is used herein for chains containing more than seven amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus. The one-letter code of amino acids used herein is commonly known in the art and can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual, 2nd, ed. Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989)

By "isolated" polypeptide or protein is intended a polypeptide or protein removed from its native environment. For example, recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the

purpose of the invention as are native or recombinant polypeptides which have been substantially purified by any suitable technique such as, for example, the single-step purification method disclosed in Smith and Johnson, Gene 67:31-40 (1988).

The amylase according to the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Protein fragments

The invention also features biologically active fragments of the polypeptides according to the invention.

Biologically active fragments of a polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein (e.g., the amino acid sequence of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18), which include fewer amino acids than the full length protein, and exhibit at least one biological activity of the corresponding full-length protein. Typically, biologically active fragments comprise a domain or motif with at least one activity of the alpha-amylase. A biologically active fragment of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the biological activities of the native form of a polypeptide of the invention.

The invention also features nucleic acid fragments which encode the

above biologically active fragments of the alpha amylase.

Fusion proteins

The proteins of the present invention or functional equivalents thereof, e.g., biologically active portions thereof, can be operatively linked to a non-alpha-amylase polypeptide (e.g., heterologous amino acid sequences) to form fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises an alpha amylase polypeptide operatively linked to a non-alpha-amylase polypeptide. In a preferred embodiment, a fusion protein comprises at least one biologically active fragment of a protein according to the invention. In this context, the term "operatively linked" is intended to indicate that the alpha amylase and the non-alpha amylase are fused in-frame to each other either to the N-terminus or C-terminus of the alpha amylase.

For example, in one embodiment, the fusion protein is a GST-fusion protein in which the alpha amylase sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant PROTEIN ACCORDING TO THE INVENTION. In another embodiment, the fusion protein comprises a protein according to the invention fused to a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian and Yeast host cells), expression and/or secretion of a protein according to the invention can be increased through use of a hetereologous signal sequence.

In another example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokarytic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

A signal sequence can be used to facilitate secretion and isolation of a protein or polypeptide of the invention. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. The signal sequence directs secretion of

the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain. Thus, for instance, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in *Gentz et al, Proc. Natl. Acad. Sci. USA 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purificaton of the fusion protein. The HA tag is another peptide useful for purification which corresponds to an epitope derived of influenza hemaglutinin protein, which has been described by Wilson *et al., Cell 37*:767 (1984), for instance.

Preferably, a chimeric or fusion protein comprising a protein according to the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g, a GST polypeptide). A nucleic acid according to the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the fusion moiety in order to express a fusion protein comprising a protein according to the invention.

Functional equivalents

The terms "functional equivalents" and "functional variants" are used interchangeably herein. Functional equivalents of the alpha amylase encoding DNA fragments described herein are isolated DNA fragments that encode a polypeptide that exhibits a particular function of the A. niger amylase as defined herein. A functional equivalent of a polypeptide according to the invention is a polypeptide that exhibits at least one function of an A. niger amylase as defined herein. Functional equivalents therefore also encompass biologically active fragments.

Functional protein or polypeptide equivalents may contain only conservative substitutions of one or more amino acids in the sequences provided in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or substitutions, insertions or deletions of non-essential amino acids. Accordingly, a non-essential amino acid is a residue that can be altered in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 without substantially altering the biological function. For example, amino acid residues that are conserved among the proteins of the present invention, are predicted to be particularly unamenable to alteration. Furthermore, amino acids conserved among the proteins according to the present invention and other amylases are not likely to be amenable to alteration.

The term "conservative substitution" is intended to mean that a substitution in which the amino acid residue is replaced with an amino acid residue having a similar side chain. These families are known in the art and include amino acids with basic side chains (e.g.lysine, arginine and hystidine), acidic side chains (e.g. aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagines, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine tryptophan, histidine).

Functional nucleic acid equivalents may typically contain silent mutations or mutations that do not alter the biological function of encoded polypeptide. Accordingly, the invention provides nucleic acid molecules encoding proteins that contain changes in amino acid residues that are not essential for a particular biological activity. Such proteins differ in amino acid sequence from SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 yet retain at least one biological activity. In one embodiment the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises a

substantially homologous amino acid sequence of at least about 40%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14; SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18.

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J.U. et al., Science 247:1306-1310 (1990) wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selects or screens to identify sequences that maintain functionality. As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require non-polar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent substitutions are described in Bowie et al, supra, and the references cited therein.

An isolated nucleic acid molecule encoding a protein homologous to the protein having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 can be created by introducing one or more nucleotide substitutions, additions or deletions into the coding nucleotide sequences having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 such that one or more amino acid substitutions, deletions or insertions are introduced into the encoded protein. Such mutations may be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

The term "functional equivalents" also encompasses orthologues of the A. niger alpha amylases provided herein. Orthologues of the A. niger alpha amylase are proteins that can be isolated from other strains or species and possess a similar or identical biological activity. Such orthologues can readily be identified as comprising an amino acid sequence that is substantially homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ

ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18.

As defined herein, the term "substantially homologous" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., with similar side chain) amino acids or nucleotides to a second amino acid or nucleotide sequence such that the first and the second amino acid or nucleotide sequences have a common domain. For example, amino acid or nucleotide sequences which contain a common domain having about 40%, preferably 65%, more preferably 70%, even more preferably 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity or more are defined herein as sufficiently identical.

Also, nucleic acids encoding other alpha amylase family members, which thus have a nucleotide sequence that differs from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12, are within the scope of the invention. Moreover, nucleic acids encoding alpha amylases from different species which thus have a nucleotide sequence which differs from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 are within the scope of the invention.

Nucleic acid molecules corresponding to variants (e.g. natural allelic variants) and homologues of the DNA according to the invention can be isolated based on their homology to the nucleic acids disclosed herein using the cDNAs disclosed herein or a suitable fragment thereof, as a hybridisation probe according to standard hybridisation techniques preferably under highly stringent hybridisation conditions.

In addition to naturally occurring allelic variants of the A niger sequences provided herein, the skilled person will recognise that changes can be introduced by mutation into the nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 thereby leading to changes in the amino acid sequence of the alpha amylase protein without substantially altering the function of the protein.

In another aspect of the invention, improved alpha amylases are provided. Improved alpha amylases are proteins wherein at least one biological activity

is improved. Such proteins may be obtained by randomly introducing mutations along all or part of the coding sequence, such as by saturation mutagenesis, and the resulting mutants can be expressed recombinantly and screened for biological activity. For instance, the art provides for standard assays for measuring the enzymatic activity of amylases and thus improved proteins may easily be selected.

In a preferred embodiment the alpha amylase has an amino acid sequence having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18. In another embodiment, the alpha amylase is substantially homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and retains at least one biological activity of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18, yet differs in amino acid sequence due to natural variation or mutagenesis as described above.

In a further preferred embodiment, the alpha amylase has an amino acid sequence encoded by an isolated nucleic acid fragment capable of hybridising to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12, preferably under highly stringent hybridisation conditions.

Accordingly, an alpha amylase according to the invention is an isolated protein which comprises an amino acid sequence at least about 40%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and retains at least one functional activity of the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18.

Functional equivalents of a protein according to the invention can also be identified e.g. by screening combinatorial libraries of mutants, e.g. truncation mutants, of the protein of the invention for amylase activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level. A variegated library of variants can be produced by, for example,

enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display). There are a variety of methods that can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide of the invention can be used to generate a variegated population of polypeptides for screening a subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations of truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

It will be apparent for the person skilled in the art that DNA sequence polymorphisms may exist that may lead to changes in the amino acid sequence of the

alpha amylase within a given population. Such genetic polymorphisms may exist in cells from different populations or within a population due to natural allelic variation. Allelic variants may also include functional equivalents.

Fragments of a polynucleotide according to the invention may also comprise polynucleotides not encoding functional polypeptides. Such polynucleotides may function as probes or primers for a PCR reaction.

Nucleic acids according to the invention irrespective of whether they encode functional or non-functional polypeptides, can be used as hybridization probes or polymerase chain reaction (PCR) primers. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having alpha amylase activity include, inter alia, (1) isolating the gene encoding the alpha amylase, or allelic variants thereof from a cDNA library e.g. from other organisms than A. niger; (2) in situ hybridization (e.g. FISH) to metaphase chromosomal spreads to provide precise chromosomal location of the alpha amylase gene as described in Verma et al., Human Chromosomes: a Manual of Basic Techniques, Pergamon Press, New York (1988); (3) Northern blot analysis for detecting expression of alpha amylase mRNA in specific tissues and/or cells and 4) probes and primers that can be used as a diagnostic tool to analyse the presence of a nucleic acid hybridisable to the alpha amylase probe in a given biological (e.g. tissue) sample.

Also encompassed by the invention is a method of obtaining a functional equivalent of an alpha amylase gene or cDNA. Such a method entails obtaining a labelled probe that includes an isolated nucleic acid which encodes all or a portion of the sequence having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or a variant thereof; screening a nucleic acid fragment library with the labelled probe under conditions that allow hybridisation of the probe to nucleic acid fragments in the library, thereby forming nucleic acid duplexes, and preparing a full-length gene sequence from the nucleic acid fragments in any labelled duplex to obtain a gene related to the alpha amylase gene.

In one embodiment, a nucleic acid according to the invention is at least 40%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to a nucleic acid sequence shown in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or the complement thereof.

In another preferred embodiment a polypeptide of the invention is at least 40%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to the amino acid sequence shown in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18.

Host cells

In another embodiment, the invention features cells, e.g., transformed host cells or recombinant host cells that contain a nucleic acid encompassed by the invention. A "transformed cell" or "recombinant cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid according to the invention. Both prokaryotic and eukaryotic cells are included, e.g., bacteria, fungi, yeast, and the like, especially preferred are cells from filamentous fungi, in particular Aspergillus niger.

A host cell can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in a specific, desired fashion. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may facilitate optimal functioning of the protein.

Various host cells have characteristic and specific mechanisms for post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems familiar to those of skill in the art of molecular biology and/or microbiology can be chosen to ensure the desired and correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product can be used. Such host cells are well known in the art.

Host cells also include, but are not limited to, mammalian cell lines such as CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and choroid plexus cell lines.

If desired, the polypeptides according to the invention can be produced by a stably-transfected cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, methods for constructing such cell lines are also publicly known, e.g., in Ausubel et al. (supra).

Antibodies

The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind alpha amylases according to the invention.

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and $F(ab')_2$ fragments) which are capable of specifically binding to a protein according to the invention. Fab and $F(ab')_2$ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl *et al., J. Nucl. Med. 24*:316-325 (1983)). Thus, these fragments are preferred.

The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the alpha amylase according to the invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of a protein according to the invention is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or alpha amylase-binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler *et al., Nature 256*:495 (1975); Kohler *et al., Eur. J. Immunol. 6*:511 (1976); Hammerling *et al.,* In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a protein according to the invention or, with a cell expressing a protein according to the invention. The splenocytes of thus immunised mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present inventoin; however, it is preferably to employ the parent myeloma cell line (SP_2O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al. (Gastro-enterology 80*:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the alpha amylase antigen. In general, the polypeptides can be coupled to a carrier protein, such as KLH, as described in Ausubel et al., supra, mixed with an adjuvant, and injected into a host mammal.

In particular, various host animals can be immunized by injection of a polypeptide of interest. Examples of suitable host animals include rabbits, mice, guinea pigs, and rats. Various adjuvants can be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), adjuvant mineral gels such as aluminum hydroxide, surface actve substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of the immunized animals.

Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD, and any subclass thereof. The hybridomas producing the mAbs of this invention can be cultivated *in vitro* or *in vivo*.

Once produced, polyclonal or monoclonal antibodies are tested for specific recognition of a protein according to the invention or functional equivalent thereof in an immunoassay, such as a Western blot or immunoprecipitation analysis using standard techniques, e.g., as described in Ausubel et al., supra. Antibodies that specifically bind to a protein according to the invention or functional equivalents thereof are useful in the invention. For example, such antibodies can be used in an immunoassay to detect a protein according to the invention in pathogenic or non-pathogenic strains of *Aspergillus* (e.g., in *Aspergillus* extracts).

Preferably, antibodies of the invention are produced using fragments of a protein according to the invention that appears likely to be antigenic, by criteria such as high frequency of charged residues. For example, such fragments may be generated by standard techniques of PCR, and then cloned into the pGEX expression vector (Ausubel et al., supra). Fusion proteins may then be expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel, et al., supra. If desired, several (e.g., two or three) fusions can be generated for each protein, and each fusion can be injected into at least two rabbits. Antisera can be raised by injections in a series, typically including at least three booster injections. Typically, the antisera are checked for their ability to immunoprecipitate a recombinant alpha amylase according to the invention or functional equivalents thereof whereas unrelated proteins may serve as a control for the specificity of the immune reaction.

Alternatively, techniques decribed for the production of single chain antibodies (U.S. Patent 4,946,778 and 4,704,692) can be adapted to produce single chain antibodies against a protein according to the invention or functional equivalents

thereof. Kits for generating and screening phage display libraries are commercially available, e.g. from Pharmacia.

Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223, 409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 20791; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246;1275-1281; Griffiths et al. (1993) *EMBO J.* 12:725-734.

Polyclonal and monoclonal antibodies that specifically bind a protein according to the invention or functional equivalents thereof can be used, for example, to detect expression of gene encoding a protein according to the invention or a functional equivalent thereof e.g. in another strain of *Aspergillus*. For example, a protein according to the invention can be readily detected in conventional immunoassays of *Aspergillus* cells or extracts. Examples of suitable assays include, without limitation, Western blotting, ELISAs, radioimmune assays, and the like.

By "specifically binds" is meant that an antibody recognizes and binds a particular antigen, e.g., a protein according to the invention polypeptide, but does not substantially recognize and bind other unrelated molecules in a sample.

Antibodies can be purified, for example, by affinity chromatography methods in which the polypeptide antigen is immobilized on a resin.

An antibody directed against a polypeptide of the invention (e.g., monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. The antibodies can also be used diagnostically to monitor protein levels in cells or tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen or in the diagnosis of Aspergillosis..

Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase,

alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive materials include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, e.g. hydrophilic regions. Hydrophobicity plots of the proteins of the invention can be used to identify hydrophilic regions..

The antigenic peptide of a protein of the invention comprises at least 7 (preferably 10, 15, 20, or 30) contiguous amino acid residues of an amino acid sequense selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein.

Preferred epitopes encompassed by the antigenic peptide are regions of a protein according to the invention that are located on the surface of the protein, e.g., hydrophilic regions, hydrophobic regions, alpha regions, beta regions, coil regions, turn regions and flexible regions.

Immunoassays

Qualitative or quantitative determination of a polypeptide according to the present invention in a biological sample can occur using any art-known method. Antibody-based techniques provide special advantages for assaying specific polypeptide levels in a biological sample.

In these, the specific recognition is provided by the primary antibody (polyclonal or monoclonal) but the secondary detection system can utilize fluorescent, enzyme, or other conjugated secondary antibodies. As a result, an immunocomplex is obtained.

Accordingly, the invention provides a method for diagnosing whether a certain organism is infected with Aspergillus comprising the steps of:

- Isolating a biological sample from said organism suspected to be infected with Aspergillus,
- reacting said biological sample with an antibody according to the invention,

- determining whether immunecomplexes are formed.

Tissues can also be extracted, e.g., with urea and neutral detergent, for the liberation of protein for Western-blot or dot/slot assay. This technique can also be applied to body fluids.

Other antibody-based methods useful for detecting a protein according to the invention, include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). For example, - specific monoclonal antibodies against a protein according to the invention can be used both as an immunoabsorbent and as an enzyme-labeled probe to detect and quantify a protein according to the invention. The amount of specific protein present in the sample can be calculated by reference to the amount present in a standard preparation using a linear regression computer algorithm. In another ELISA assay, two distinct specific monoclonal antibodies can be used to detect a protein according to the invention in a biological fluid. In this assay, one of the antibodies is used as the immuno-absorbent and the other as the enzyme-labeled probe.

The above techniques may be conducted essentially as a "one-step" or "two-step" assay. The "one-step" assay involves contacting a protein according to the invention with immobilized antibody and, without washing, contacting the mixture with the labeled antibody. The "two-step" assay involves washing before contacting the mixture with the labeled antibody. Other conventional methods may also be employed as suitable. It is usually desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with the component and readily removed from the sample.

Suitable enzyme labels include, for example, those from the oxidase group, which catalyze the production of hydrogen peroxide by reacting with substrate. Activity of an oxidase label may be assayed by measuring the concentration of hydrogen peroxide formed by the enzyme-labelled antibody/substrate reaction.

Besides enzymes, other suitable labels include radioisotopes, such as iodine (^{125}I, ^{12l}I), carbon (^{14}C), sulphur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

Specific binding of a test compound to a protein according to the invention can be detected, for example, in vitro by reversibly or irreversibly immobilizing a protein according to the invention polypeptide on a substrate, e.g., the surface of a well of a 96-well polystyrene microtitre plate. Methods for immobilizing polypeptides

and other small molecules are well known in the art. For example, the microtitre plates can be coated with a protein according to the invention by adding the polypeptide in a solution (typically, at a concentration of 0.05 to 1 mg/ml in a volume of 1-100 ul) to each well, and incubating the plates at room temperature to 37 ^{o}C for 0.1 to 36 hours. Polypeptides that are not bound to the plate can be removed by shaking the excess solution from the plate, and then washing the plate (once or repeatedly) with water or a buffer. Typically, the polypeptide is contained in water or a buffer. The plate is then washed with a buffer that lacks the bound polypeptide. To block the free protein-binding sites on the plates, the plates are blocked with a protein that is unrelated to the bound polypeptide. For example, 300 ul of bovine serum albumin (BSA) at a concentration of 2 mg/ml in Tris-HCl is suitable. Suitable substrates include those substrates that contain a defined cross-linking chemistry (e.g., plastic substrates, such as polystyrene, styrene, or polypropylene substrates from Corning Costar Corp. (Cambridge, MA), for example) . If desired, a beaded particle, e.g., beaded agarose or beaded sepharose, can be used as the substrate.

Binding of the test compound to the polypeptides according to the invention can be detected by any of a variety of artknown methods. For example, a specific antibody can be used in an immunoassay. If desired, the antibody can be labeled (e.g., fluorescently or with a radioisotope) and detected directly (see, e.g., West and McMahon, J. Cell Biol. 74:264, 1977). Alternatively, a second antibody can be used for detection (e.g., a labeled antibody that binds the Fc portion of an anti-AN97 antibody). In an alternative detection method, a protein according to the invention is labelled (e.g., with a radioisotope, fluorophore, chromophore, or the like), and the label is detected. In still another method, a protein according to the invention is produced as a fusion protein with a protein that can be detected optically, e.g., green fluorescent protein (which can be detected under UV light). In an alternative method, a protein according to the invention polypeptide can be covalently attached to or fused with an enzyme having a detectable enzymatic activity, such as horse radish peroxidase, alkaline phosphatase, a-galactosidase, or glucose oxidase. Genes encoding all of these enzymes have been cloned and are readily available for use by those of skill in the art. If desired, the fusion protein can include an antigen, and such an antigen can be detected and measured with a polyclonal or monoclonal antibody using conventional methods. Suitable antigens include enzymes (e.g., horse radish peroxidase, alkaline phosphatase, and a-galactosidase) and non-enzymatic polypeptides (e.g., serum proteins, such as BSA and globulins, and milk proteins, such

as caseins).

Epitopes, antigens and immunogens.

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen, H. M. et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1984).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J. G. et al., Science 219:660-666 (1984). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer, soluble peptides, especially those containing proline residues, usually are effective. Sutcliffe et al., supra, at 661. For instance, 18 of 20 peptides designed according to these guidelines, containing 8-39 residues covering 75% of the sequence of the influenza virus hemagglutinin HAl polypeptide chain, induced antibodies that reacted with the HA1 protein or intact virus; and 12/12 peptides from the MuLV polymerase and 18/18 from the rabies glycoprotein induced antibodies that precipitated the respective proteins.

Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. Thus, a high proportion of hybridomas obtained by fusion of spleen cells from donors immunized with an antigen epitope-

bearing peptide generally secrete antibody reactive with the native protein. Sutcliffe et al., supra, at 663. The antibodies raised by antigenic epitope bearing peptides or polypeptides are useful to detect the mimicked protein, and antibodies to different peptides may be used for tracking the fate of various regions of a protein precursor which undergoes posttranslation processing. The peptides and anti-peptide antibodies may be used in a variety of qualitative or quantitative assays for the mimicked protein, for instance in competition assays since it has been shown that even short peptides (e.g., about 9 amino acids) can bind and displace the larger peptides in immunoprecipitation assays. See, for instance, Wilson, I.A. et al., Cell 37:767-778 at 777 (1984). The anti-peptide antibodies of the invention also are useful for purification of the mimicked protein, for instance, by adsorption chromatography using methods well known in the art.

Antigenic epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (i.e., the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); and sequences containing proline residues are particularly preferred.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies.

Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis. For instance, Houghten has described a simple method for synthesis of large numbers of peptides, such as 10-20 mg of 248 different

13 residue peptides representing single amino acid variants of a segment of the HAI polypeptide which were prepared and characterized (by ELISA-type binding studies) in less than four weeks. Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten et al. (1986). In this procedure the individual resins for the solid-phase synthesis of various peptides are contained in separate solvent-permeable packets, enabling the optimal use of the many identical repetitive steps involved in solid-phase methods.

A completely manual procedure allows 500-1000 or more syntheses to be conducted simultaneously. Houghten et al., supra, at 5134.

Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art. See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F.J. et al., J. Gen. Virol. 66:2347-2354 (1985).

Generally, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as keyhole limpet hemocyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine may be coupled to carrier using a linker such as maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde.

Animals such as rabbits, rats and mice are immunized with either free or carriercoupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 ug peptide or carrier protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. For instance, Geysen et al., 1984, supra, discloses a procedure for rapid concurrent synthesis on solid supports of hundreds of peptides of sufficient purity to react in an enzyme-linked

immunosorbent assay. Interaction of synthesized peptides with antibodies is then easily detected without removing them from the support. In this manner a peptide bearing an immunogenic epitope of a desired protein may be identified routinely by one of ordinary skill in the art. For instance, the immunologically important epitope in the coat protein of foot-and-mouth disease virus was located by Geysen et al. with a resolution of seven amino acids by synthesis of an overlapping set of all 208 possible hexapeptides covering the entire 213 amino acid sequence of the protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope were synthesized, and the particular amino acids conferring specificity for the reaction with antibody were determined. Thus, peptide analogs of the epitope-bearing peptides of the invention can be made routinely by this method. U.S. Patent No. 4,708,781 to Geysen (1987) further describes this method of identifying a peptide bearing an immunogenic epitope of a desired protein.

Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. et al. (1996) on Peralkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Use of alpha amylases in industrial processes

The invention also relates to the use of a protein according to the invention in a selected number of industrial and pharmaceutical processes. Despite the long term experience obtained with these processes, the amylase according to the invention features a number of significant advantages over the enzymes currently used. Depending on the specific application, these advantages can include aspects like lower production costs, higher specificity towards the substrate, less antigenic, less

undesirable side activities, higher yields when produced in a suitable microorganism, more suitable pH and temperature ranges, better tastes of the final product as well as food grade and kosher aspects.

An important aspect of the amylases according to the invention is that they cover a whole range of pH and temperature optima which are ideally suited for a variety of applications. For example many large scale processes benefit from relatively high processing temperatures of 50 degrees C or higher, e.g. to control the risks of microbial infections. Several alpha amylases according to the invention comply with this demand but at the same time they are not that heat stable that they resist attempts to inactivate the enzyme by an additional heat treatment. The latter feature allows production routes that yield final products free of residual enzyme activity. Similarly many feed and food products have slightly acidic pH values so that amylases with acidic or near neutral pH optima are preferred for their processing. An alpha amylase according to the invention complies with this requirement as well.

SEQUENCE LISTING

<110> DSM NV

<120> NOVEL AMYLASES AND USES THEREOF

<130> 20438WO

<160> 18

<170> PatentIn version 3.1

<210> 1
<211> 3001
<212> DNA
<213> Aspergillus niger

<400> 1
ggttataatt caaaattcaa cttccacctt tgtttcaccg gcggccacgg cattcctgca 60

tgactaacgt tctgtaaatg gacccgataa cacccagcac gttgcagcag agaaggtact 120

ctctcacacg cactgctctt tatagttgcc gagacggccg ccgaggagaa aaccgccggc 180

ctgtggccac tattcgctgg aaggaaccct gccagtcgaa cacacccgcc cgtgatcgcc 240

aggggccgat ggatttcccc ccgcatcctt gtcggttcat gagtgaagac tttaaatccc 300

atctagctga cggtcgggta catcaataac tggcagccta gtttccaaga cacggagaag 360

catgtaatcg ctatttatag aatgctggga tcggacccgt cgaatggtct tccgatggga 420

agtgacaact cacattgtca tgttggcctt actcaatcca acgggatctg acctgctttg 480

gctaacctag tataaatcag catgtctctc ctttgataca tcggatcgtt cctcaaatat 540

agttatatct tcgaaaaatt gacaagaagg atgacaatct ttctgtttct ggccattttc 600

gtggctacag ctctggcagc cacgcctgca gaatggcgct cccagtcgat atatttcctg 660

ctcaccgatc gctttgcgcg aacggataat tctaccactg cttcttgtga cttgagcgct 720

cgggttagtc acagcatgtt ctagaatctc caattgattc gctgacagat ctagcaatat 780

tgcggtggat cctggcaggg catcatcaat caggtcggtc cgtccatcgt tgcagcacta 840

tctacatcaa cgtttgtttg gcaaattaac atccattagc tggactatat tcaaggaatg 900

ggctttacag cgatctggat cacacccgta actgcacaga tcccccaaga tactggttac 960

ggacaggcat atcacggata ctggcagcag gacgcgtgag atgctacctc tatcgcccgg 1020

atgaatgtat atccttctta ccatgcagac agttatgccc tgaactccca ttatggtacg 1080

gcagacgatc tcaaagctct ggcttcagct cttcactcac ggggcatgta tctcatggtg 1140

gacgttgttg ccaatcacat ggtatgttct tagcctccca cgggacctta gctttatatc 1200

tgacagcgat agggccacaa tggtacgggg agctctgtgg actacagtgt ttataggcca 1260

tttaattcgc aaaagtactt tcacaacctc tgttggatct ctgattacaa taaccagaca 1320

aacgttgaag actgctggct aggcgataac accgttgcct tgccggatct tgatactacc 1380

agtacggagg tgaagaatat gtggtatgac tgggtcgagt ctctcgtctc taactactcc 1440

ggtaatccta cctttacttc gctattttct gcctcttatg agacaaagac taacaaatat 1500

caagtcgacg gcctccgcgt agacacagtc aagaacgtac agaagaactt ctggcccggc 1560

tacaacaatg cttcaggcgt gtactgtatt ggagaagtct tcgatgggga cgcctcatac 1620

acctgtcctt atcaggaaga cttggacgga gtccttaatt accccatgta agccctacat 1680

ttaaccccat tgaatgcttg ccaacgactt gcaataggta ctatccactc ctccgggctt 1740

tcgaatccac caacggcagt atcagcgacc tctataacat gatcaacacc gtgaaatcca 1800

cctgcagaga ttctacgctt ctagggacct tcgtcgaaaa ccacgataac ccacgctttg 1860

ccaagtaaga atatcctctc cgagttcacc attacaaaca caagagctca cctcagaagc 1920

tacacaagcg acatgtccct agccaaaaat gccgcaacat tcactatcct ggctgacggc 1980

attcccatca tatacgccgg tcaggaacag cactatagcg gcggtaatga cccctacaac 2040

cgcgaagcga cctggctttc aggctacaag accaccagcg agctctacac ccatatcgcc 2100

gcatcgaaca agattcgcac ccacgctata aaacaggata ccggatatct cacctacaaa 2160

gtaatcttca ttcgagtcca tgtgtggtac aatctatctg actagaacct attctagaac 2220

taccccatct accaagacac ctcgaccctt gccatgcgca aaggctacaa tggcacccaa 2280

actatcacgg tcctttctaa ccttggcgcc tcggggtcct catacacact ctccctccca 2340

ggaacaggct acacagccgg ccaaaagatt actgaaatct atacctgcac gaatctaaca 2400

gtcaactcaa atggctcggt gccagtaccc atgaagagcg ggttaccgcg gatcctctat 2460

cctgcagata agttggttaa tggaagctca ttttgcagtt agttccctgc ctatatgttt 2520

caataagccg tatttgcatg cgcgtccatc gcattgattc tttatggatt gatcaactga 2580

tacattcgca caggctctgc gatcactgtt ttcaaggatg caggaggagg cgtgttgttc 2640

tttagctaca ctgtgtaagt gtatcagatc atatcatggc cactagacga tgacccctgt 2700

gatcgtgatc cagatgctga cttcaatttg tagaattttc gctcaggtgc ttattgctat 2760

aatgacgtga ttaggatgtc tcagatgaca gtatatatta ttaggtgacg tgatgattct 2820

atgtgtaatg gattcttcaa gatcatttta tattattagt atccacagta ttcatttgcc 2880

tgcaaagtaa gtatggataa aatcactacc accgacgtag tagttgtacc tagtccttag 2940

tgtctcaaag ccacagtcac tacaacctca cgaaacgacg gagtaaactt atttctattg 3000

a 3001

<210> 2

<211> 2511

<212> DNA

<213> Aspergillus niger

<400> 2

catgaacgct ttcccagccg aaatgtggga agttaatttg ccttgtccac ttcgccctcc 60

agcgcagccc tgccctcacg acaatcagtc aatctcccgt gacaatccca ggggatcttc 120

gtgttctaac gctgctaccc gttctcgaat gttcgggtga actggagtac tagagtagat 180

taccatcgag cgttacaaat ttagatggcg tctttatttt tggtccgcgt tgagggtcgt 240

tgaaagtgat ggatcagggc ttcgggcacg cgatgcgaac gagggtgggg gagagactac 300

ctatttatgg tacctccggc ttctaccggc cttgtctgca gccccccggt ctggcgctct 360

tatctctctt gctaagttgc tatagctatc gggggtccgc ccccccctgaa tcgcgttgcc 420

ttgacttgtc attctttctt ttccccgtta ttgcaacggg aacagcgcac ccggcctcaa 480

acatcatgct cagcggcccg tttggctgaa tcgcgtgaaa ttgaaacgct taaaacacat 540

tcgttagagt aagctgttgc ttgtcataat tcattcggtc acccaagccc acccacggga 600

agcagaacct ccacaagggg gcacagaggg ttgatcgagg gtagtactgc caatgatgtt 660

tcgaaaatcc gcttccctcc tgggccaacg gctcatggcc gtttgtctcc tgtgctggtg 720

cgtttcgcta gcgaccgccg caagcacaga agaatggaag acgcgatcca tctaccagac 780

gatgacggat cggtttgccc tcaccaacgg ttcgacgacc gcaccatgca atactacggt 840

agccaactac tgtggcgggt cttggcaggg gacgatcgat aagctggact acatccaagg 900

catggggttc gatgcgatca tgatctctcc cgtaattaaa aacattgcgg ggcgatctaa 960

ggacggtgag gcctaccatg gatactggcc tctggatcta tacgagatca attctcattt 1020

cggaacccgg gaggaactgt tgaagctgag cgaggagatc cacgcacgcg gcatgtactt 1080

gctgctggac gtcgtcatca ataatatggc ttacatgacg gacggcgagg atcctgcgac 1140

gaccattgac tacaatgttt tccctcagtt caatggatct tcatacttcc acccctactg 1200

tcttattacg aactggaata actatacgga cgcgcagtgg tgtcagactg gtgacaatta 1260

tactgcactc ccagatctgt acacagaaca caccgcggtg cagaacatct taatggactg 1320

gagtaaatcg gtcatcagca attactccgt cgacgggcta cgaatcgacg ctgccaagtc 1380

cctcactccc agctttctgc ctacatatgc gagtaccgtg ggcggattca tgaccggcga 1440

agtcatggat tcgaacgcca ccaacgtgtg caaatatcag acggattacc tgccaagtct 1500

tccaaactac cccctctatt actccatgat cacggccttc ctcaacggcg agcctgcgac 1560

cttgctcgag gaaatcgcga cgatcaatga tctttgccct gatacgttcg caatggtcaa 1620

cttcatcgaa gatcaagacg ttgaccgatg ggcctacatg aatgacgaca tcatgctagc 1680

taaaactgca ctgaccttca tgatgctcta cgacggtatt cccttggtct accagggtct 1740

ggagcaggcc attgcctatt ccaaccgagc ggccttgtgg ttgacagatt tcgacaccaa 1800

tgcgacgctt tataaacaca tcaagaaact caatgccatc cgcaaacatg ccattaacct 1860

tgattccagc tacatcagtt cgaaaacata tcccatctat caaggaggta gcgagttggc 1920

tttctggaag ggcaacaatg gacgccaagt catcatggtt ttatctacgg cgggctcgaa 1980

tggttcagct tatactctga cactacccgt aagctatggg gctagtgagg tggtcacgga 2040

agtgctgaac tgcgtcaatt atacggtcaa cacttacagt caattggtcg tggacatgga 2100

taagggcgaa cctcgggtct tctttcccgc ttcgatgatg ccggggagcg ggctgtgtgg 2160

ctacaatact tctaatgtga cgtattctga gctgagactt gctgctgtgg gatcttcgtc 2220

gtctgctggt agccactctg tcataccgtc tgcttttgct tcgcttttca tggcgatagt 2280

tgcatttttg gcattccgga tataagttca tatttttttct actcagtttt ataccagcaa 2340

ataaaacgat gtcctatgaa tatacacatt tcaaaagtgg agacaaaatt atatctttat 2400

atccctagca cctccctttc cgtatccttt cctgcaataa agacctattc agcattggga 2460

gcagaattac gaagcacagt ctcctcagtt cgcatacccg cccacctcca a 2511

<210> 3

<211> 2501

<212> DNA

<213> Aspergillus niger

<400> 3
ccccagcatt ccacttcttc tcgcttcccc tcttctctct cctctccctc tcctctcctc 60
tcctttccct tcccttccct tccttcctcc tcactcattc tttcgctctc gtcacatcga 120
ccaatgcatc cgtcactgaa gtgacccccc cttcctgggc tgttcccttg ccattcattt 180
ttttctttgg atctgaccat tgccattctt tctctacaat ctggtcccac ttctttcgtt 240
cactctttca tttcgtgcct agcccgtatc cttcctttta gcgtaaccgg ctatctccca 300
gcattcgaca gtcaccttat cgtatctcca agtcgggccc cgcaattgct cggcaggcac 360
atgtacgcaa cccttgatca ttaagataat tccaatctcc cagtgactgt agccatggtc 420
tcaatgtcgg ccctgcggca cgggcttggg gtcctctatc ttgctagctg gctggggtcg 480
tcgctggctg ccagcaccga gcaatggaag tcccggtcaa tctatcagac catgacggat 540
cggttcgcgc gcactgatgg ctcaaccacc tccccctgca acaccacgga gggtctgtac 600
tgtggtggta cctggcgcgg catgatcaac catctggatt acatccaggg gatgggcttt 660
gatgccgtca tgatctcccc tatcatcgag aacgtcgaag gtcgcgtgga gtacggagaa 720
gcctatcacg gctattggcc cgtggatctg tactccctca actcgcactt tggaacccac 780
caggatctac tggacctgag tgatgccctg catgctcgcg acatgtatct gatgatggac 840
acagtcatca acaacatggc ctacatcacc aatggctccg accccgccac ccacatcgac 900
tactcaaccc tgacccccct caatagctcg tcttactacc atccctactg taagatcacc 960
gactggaata acttcaccaa tgcccagctg tgccaaaccg gtgacaacat cgtggccttg 1020
cccgatctgt ataccgagca tgccgaggtg caagaaaccc tgagcaactg ggccaaggaa 1080
gtcatctcca cctattccat cgacggcctc cgcatcgatg ccgcaaaaca tgtgaacccc 1140
ggcttcctga agaatttcgg cgatgcgctt gatatcttca tgaccggcga agtgctgcag 1200
caagaggtca gcacaatctg cgattatcag aacaactaca tcggcagtct cccgaattac 1260
cccgtctact acgccatgtt gaaggccttt accttgggca acaccagcgc cctggccact 1320
caggtccagt cgatgaagaa ttcctgtaat gatgtgaccg ccttgtcgtc attctccgaa 1380
aaccacgatg tcgcacgatt tgccagcatg acccatgaca tggctgtatg aactctgctt 1440
gactttacag acccctttcc tagctaacat tgtccctact agctcgcgaa gaatattctc 1500
acattcactc ttctcttcga cggtgttccc atgatctatc aaggtcaaga gcagcatctc 1560
gacggccctg gcagtccgga aaaccgagaa gcgatatggc tctctgagta caacaccgac 1620
gccgagctct acaagttgat cggcaagttg aacgccatcc ggaagcacgc ctaccgactc 1680

gacaatcact accccgatgt cgagacgtac cccatcttcg aaggtggcag tgaactggga 1740

ttccgcaagg gaatcgaagg tcgacaggtg gtcatgcttc tgtccacgca aggcacgaac 1800

agcagcgctt acaacctctc gatgccggtg agtttcacgg gtggcacagt agtgaccgag 1860

attctgaact gcgtcaacta cacggtcaac acccagagcg aactcgtggt tccgatggat 1920

aaaggagagc cgcgcgtgtt cttcccggcg gatctgatgc ccggcagtgg gctgtgcgga 1980

cttcctgtcg cgaatgtcac ctatgctgcc ttgaggacgc agggtgcagc ggcggcggag 2040

gctgccttgt cgctaggtat caagactgat gcagcttcta gtgctttgct ttctctgggg 2100

ctgtctgtgg tggcgggtct gattgtgggg atgtggtaat atgctgatag ctacagatgg 2160

gttgagatat attatcatgt ttcagtgttc tttgtatgta tttgcagtaa cgataccctt 2220

acctagtata ttgtatatat tcacttcgct atatccacta ctactcgcag atactagccg 2280

catctcgctt tattccacac tagttactag cacttcagca ctaagcaata aacagaacga 2340

agacaaatta ctagtccagt atatatataa atagagatca cctgtaaatt cttcagaaaa 2400

cccaaacggg tacctgcatc caacagctgt tgacgatcat ggctggctga tcctctggag 2460

gcaaatccgt tgggcctcgg tagctttccg aggaaaccgc c 2501

<210> 4

<211> 3080

<212> DNA

<213> Aspergillus niger

<400> 4
catctcgctt tattccacac tagttactag cacttcagca ctaagcaata aacagaacga 60

agacaaatta ctagtccagt atatatataa atagagatca cctgtaaatt cttcagaaaa 120

cccaaacggg tacctgcatc caacagctgt tgacgatcat ggctggctga tcctctggag 180

gcaaatccgt tgggcctcgg tagctttccg aggaaaccgc ccgagctaaa aatactcgct 240

catcagccag acaggcccac tttaaggacc cattgaaaat tgaaatgaag tcatgcacac 300

cgtgttaact agctaactag ctagttcttc ttagtgtcaa gacagcagct attctaacgt 360

tcttccgtgc ttacaccaca gcccagctcc gtcgctgcat tcgactcttc accttccatc 420

catcatccat ccattctaca tggacacaga cagtaaacct agctaatccc tcctccgcac 480

cgcagtaaaa aaatcctcac ctcccgaacc accaacgacc tacaaacata acccaagaaa 540

aaaagaatac aagcaaaatg ttcagtttcc tcccctgctt caaaactcgc cgacaacgca 600
ccaaatccca aactcaatcc caaaagcaaa tagagggtta gccaaacacc cccaactacc 660
cataccaagt caattcaatt caaatcatag taacaaattc aacctcccca atagaaaaag 720
ccaacgccat agaatccctc ccctcatggc actcccccac ggaaaacacc ctcctatttc 780
aagccttcga atggcacgtc cccgcaaccc ccaacaccgc ggacaagcgc agccactggc 840
gccggctcca gcatgcctta ccagctatcc actccctcgg tgtaaccagc atctggatac 900
cgcctggatg caagggcatg gacacaaatg gaaatgggta tgatatctac gatttatatg 960
atctaggcga attcgaccag aaaggagcag tgcgaacgaa atggggcacg cgcggagaat 1020
tagaggatct tgttcgagat gcaaatgctc tgggcgtagg tgtattatgg gatgcggtgt 1080
tgaatcataa ggctggggcg gatagcgtgg agaggttcga gggtgtgagg gttgatgatg 1140
atcgtatata ccccttccc atctctccgt tcttctatgt atagtgtgta tattgacaaa 1200
tgcttcacag gacgagatat cgaagatggt aatccccaac aaatctccgg ctggacgtcc 1260
ttcaccttcc ccggccgcgg caccacatac agccccccttc aataccactg gcagcacttc 1320
agcggcgtcg actgggacga cgcccaacaa cgcaaagcaa tctacaagat cctcgaccct 1380
tcccgccccg acaagaactg ggcccaggac gtcggcacag acgagaacgg caactatgac 1440
tacctgatgt tcgcggacct ggacttctca cacccggagg tccgggagga tgtactgcgc 1500
tgggggaagt ggattatgtc tgtgttgccg ttgagcggaa tgcggttaga tgcggcgaag 1560
catttctcga cggcatttca gagggacttt attgattgtg tgcggcagga ggccggggac 1620
aggaaggtat ttgtgattgg ggagtattgg agtggggagt taagggcgct gcttaggtat 1680
ctggaggaga tggagtatcg ggtggcggcg gtggatgtgc cgctggtaga gaggttctcg 1740
aggttgtcga gggtgaagag ggcggatttg aggggagtgt tgaggggggac gttggtggag 1800
agcaggccgg ggaatgcgtt ggtgagttta tcccagtgat ccaattcggg ggtggtgagg 1860
ctgatgtgtg tgtggtgtag acattcgtta cgaatcatga tacggtgagt gaaagccctt 1920
gtgaagtgaa ggtatatgga tgagtgctga atgatatcac aagcaaccgg gacaaatgct 1980
cgaggtgagt tgttaacttt atcatgtgtc tgaggatgag caacgagcta acatacaaca 2040
gaccatcgtc gagccctcct tcaaacccct agcctacgcc ctaatcctcc tccgccaagg 2100
aggccatcca tgtgtcttct acggcgacct ctacggcacc tgcgacggcg accacccgcc 2160
aactcccgcc tgcgagggcc aacttccgaa tctgatgcgt gcccgcaagc tgtacgccta 2220
cggtgagcag gaggactact tcgaccagcc caactgtatt ggtaagtcaa gccctatcca 2280

gtctgcgaca tagtctgata aatgagaagg attcatccgc tacggcaacg ccgcccaccc 2340

ctccggactg gcctgcgtca tgagcaacgg cggaccagcc acgaagcgca tgtacgtggg 2400

tcggaaacat gccggcgaga agtggacgga tctgctgcag cgaggcggtg atcacccgtc 2460

tgtcacggtg attgatgaaa tggggtacgg ggagtttccg gttcagagta tgagggtgag 2520

tgtctgggtg gatagtgcgg cagatggccg agagggtgtc ggggcagaat tgtacgtcta 2580

taccctattc cacttattgc tcttggtagt cgtgggggtg gatgtttaaa actgatccag 2640

gggatgatga tgatgctaat attggaagtt tagtgacgtc gatatctacg gcattcaggc 2700

actataacta acccacccgg tacttactca ttattgagaa aataatggtg tccatatata 2760

tctctaagac acgatgtata ttaatacaat tcacaatgta actatttccc agactagttt 2820

tactatagta agacaataca acaatactac agagtaataa tcaatgtatc cacacaaact 2880

agtctactcc atctattcta caactaactc caagacacac aatcacacac agtaagacac 2940

aatctcaccc aaaaataatc cctccattca tccaagcgag cgagaaaccc aaaacccgaa 3000

agctaaaaga aaagaagaaa aagattctgg accctcttta cccaatccag taacccattc 3060

ggtgctgccc gacactggta 3080

<210> 5

<211> 3010

<212> DNA

<213> Aspergillus niger

<400> 5

aaataaccat tccatcctca ccatcaagtg acgaaacctg ccatggccgc caagccgtgc 60

tgcccactga ttggtgcagt ttctcctttc ggtcaacggt aagccgacgg agaacttctg 120

cttggttcat aacgatttgt gcggtttcat cccctcaggc atcaacattt tatcctcgga 180

attggtgtct agcgttgcga cgcgattgat gcttgaccta catacttaga ccctagacca 240

gtatacgtgg tggcatctgc aaggaacgct ctagaaatca taggagaacc tcctagggat 300

atgttttgcg cccggaaggc acagcattca gccccatcta ctatttaaag tatcaacaca 360

ctaaaacagg aggctttagt ggctgttttt ctgatcgacc atcatacctt aaatcatacc 420

agacagctat caccaatatg acgtgttttg cagtatatac atttgataca tctaagacct 480

gaaccacggt ggaggctctc tttttcatca ccgaatcaat ctttctcagc ttcacactcg 540
aaataggata cttctccttc cttttcaccc ttgttgatac cgtttctgaa agtagaagta 600
cggcattttc tccatgttat ctttcctcct atggtgccat ccaaagaaga ggaaggagag 660
acaactatgg aagcagattg aaggttagcc tccgtaatca tcaaggtgca ttttattgca 720
agttgaatcc ccgcttacgt gtccaccaac agaagaagct gagcacttgg accagcttcc 780
atcatgggat gctcccgaca acactttaat gctgcaagcc tttgaatggc atgtgccagc 840
tgaccaaggt cactggcgtc gtcttcacca ggccttacca aacttcaagg cgattggggt 900
agacaacata tggatcccgc ccgggtgtaa agctatgaat ccatctggta acggatatga 960
tatttacgat ttgtatgact tgggagaatt cgagcaaaag gggtctcgag ctactaaatg 1020
gggtaccaaa gaggagctcc agtccttggt agctgcagcc caagatttcg gtataggtat 1080
ttactgggac gctgtcctca atcacaaggc gggagcagat tacgcagagc gctttcaagc 1140
tgtcagggtt gacccacagg gtatgccatg gccttattcg actagtttcg ggtgcacctg 1200
ttaacttgga attgtctaga gcgtaatatg aaaatcgccc ctgcagagga aattgaaggc 1260
tgggtgggat tcaacttctc tgggcgtggc aaccactata gttcgatgaa gtacaacaaa 1320
aaccacttca gcggtatcga ctgggaccag tcgcgtcaaa aatgcggagt ctacaagatc 1380
caaggacatg aatgggcgaa cgacgtcgcc aatgagaacg gaaattacga ctatctcatg 1440
ttcgccaact tggactactc caacgcagaa gtacgacgcg atgttctgaa atgggccgag 1500
tggctcaatg ctcaattgcc tctaagcggc atgaggttgg atgcggtcaa acattactcg 1560
gctggttttc agaaagagct tattgatcat cttcgaacta ttgctgggcc agactatttc 1620
atagtgggcg agtactggaa aggcgagacc aagccgttag ttgactacct gaagcagatg 1680
gactacaagc tatcattgtt cgattccgct ctggttgggc ggttctcaag catttcacag 1740
acaccagggg cggatcttcg caacattttc tataatacat tggtccaatt gtacccagat 1800
cattctgtcg taagttgatt gtcattttgg tgagatccct tactgaccac taaaagactt 1860
tcgttgcaaa ccatgatact gtaggtgcaa tatgcccttt ggcctcattc tactaacata 1920
gatccaaagc aaccaggtca atccctcgaa gtacgcacac cgcccacttt gattttacag 1980
atctaacaga tgttcaggcg ccagtaacat cattcttcaa acccctcgcg tacgccctta 2040
tccttctccg tgaccaaggg caaccatgca tattctacgg agacctttac ggcctccaag 2100
ccgatgtcaa agatccaatg acaccgtctt gcaggggcaa gctgtccatc ctcacccgag 2160
ctcgaaagct ctacgcatat ggcctgcaac gagattattt tgacaagccg aactgcatcg 2220

gttagtgcct gctttcgtta tcctctaccg cacactgaca ggagtaggtt ttgtccgcta 2280

tggtaaccgt cggcatccct ctggtcttgc atgcgtgatg agcaatgcgg gtccgtcgag 2340

gaagcggatg tatgttggtc gacgacacgc caagcaaaca tggaccgata tcctgcagtg 2400

gtgtgatcag actgttgtca ttgatgccaa gggatatgga gagtttccgg ttagtgcgat 2460

gagcgtgagt gtatgggtga actccgaggc tgaggggaga gatagcctct cacatcattt 2520

gtatgtccct gctcacagtc ttagtacgga tgatttgctg acgagtgctg aaattagtga 2580

tgagaacata tacaaactag cttaatcgcc ttgggtatct gtgatgacta ctagacaatc 2640

cccaaagaaa cataatacac gatcatctat acaacaacct ctctaaccaa ccaatcaaac 2700

caacaacggc gtatcaaagc acccaaaccc cttgaccaca ctcatctcat ccacatccgc 2760

aaacttcgtc tttaacgacc agcccacagt ttcatttctc aaaatctccc ccttcttcca 2820

aatacacagc ccaaacaccc ccacttccga aatcacatcc gattcctcca ccccttcccg 2880

cgacatcagc acattcccca caacaggcga cctaatcctc tccatcaaca catacctacc 2940

ccattcggat tccgaactca ttcccaacag aaaccccggg atatcctccc cgtagatatt 3000

atgccccccg 3010

<210> 6

<211> 3001

<212> DNA

<213> Aspergillus niger

<400> 6

tctgttggtt cggggtccag aagtaccaat acccaattat tcccacttta cccaatggtc 60

gaataataca caatacccat tacacaattc ccgatgccga tgtcacgtaa gtttccatac 120

ggtccggata catcggtatc atcttcgctg aagtgttgtt cactggtcgg tatcagcctt 180

cttgccctga acagacacaa aggatgtatg tgtccgcttc gtatatgaac atgctcaaca 240

ggaggttacg caggtcataa ccatcaagac atttttctc acctctttcc agctaataat 300

cccaccattg aactactctc cttgccttct ttcatctgcg catgcaaacg ccaccaacca 360

tggacgacgg gtggtggatc atctcgttgc tggtggtaac tctgggaatc cccacggtta 420

atgcggcttc cagagaccaa tggatcggtc gctctatcta tcaaatcgtg accgatcgct 480

ttgctcgatc ggataactca accaccgctg cttgtgatgc agcactagga aactactgcg 540

gaggctcttt ccagggcatc atcaataagc tggactacat ccaggagctc gggtttgatg 600

cggtaggttt accaccatca tcctgcctct cgctctctgc cactctttct catacctgaa 660

acaccagata tggatctctc ccgcacaaag ccaaatttcc gcccgaacag cagatctctc 720

aggtatggca ctttcaaaat cttcacaaac catctgcaac tctgtacatt gttatataca 780

cgatatacta tgtagcttgg taggtaagaa attcaacatg ttaacggttc atagcatacc 840

atggatattg gcccaatgat ctgtattcca tcaactctca ttttggcact cccaaggagc 900

tggaagcctt gtcctctgcc ctgcacgatc gtggcatggt gagtcgctct gccattttcc 960

ctaaagatga cactgaccgc gcaaaaccca gtacttgatg cttgacatcg tggttggtga 1020

tatggcctgg gcgggaaatc acagcaccgt cgattacagc aactttaatc ctttcaatga 1080

tcagaagttt ttccatgatt tcaagctcct ctccagtgac cccacaaatg aaacttgtgt 1140

tctggatgta agtctgtatt gtgaatcatg acccgtggga ttttgctaac cccatatgaa 1200

gtgctggatg ggagacaccg tcgtatcact tcccgatctg cgaaatgaag accaacaagt 1260

tcagaatatt cttggtacct ggatttcggg gttggtttcg aactactcaa gtaagacttc 1320

ttccttggcc gttatggaat attgattgac ctgtttctgt tttgtagttg acggactgcg 1380

tattgacagc gtcttaaata tcgccccgga cttcttctcc aacttcacca agtcatcagg 1440

ggttttcact gtcggtgaag gtgccacggc cgatgcggct gatgtttgcc ctctgcagcc 1500

aagtttaaat gggcttttga attatccatt gtaagtaccg catcgtggtg agtaaagtca 1560

gtcgtgttga ctttctcagg tactatattc ttaccgacgc cttcaacacg accaacggga 1620

acctgagcac cattaccgag tccataagct acaccaaagg acagtgcgag gtagggcatg 1680

ccttcgaatt ggattataag ggcaaactaa tcatatacag gatgtcttgg ctctggggac 1740

gtttactgca aaccaggatg ttccgcgctt cggttcatac acgtcggata tttcggtaag 1800

ttcagtcggt ggtcctatct tggtactgaa caatgtttga ctttattgct tgtgagcagc 1860

tggcgcgcaa tatcctgacc agcagtatgc tgacggacgg catccccatc cgtatgttcc 1920

gatccccgcc tcccccatag taagcaagat tcctcagaca ccactgataa tcaaaataca 1980

gtctattacg gcgaggagca gcatttgaca ggatcctaca atcctgtcaa ccgtgaggct 2040

ctgtggctga ccaactactc gatgcgctcg acctccctcc ccaccctcgt ccaatccctg 2100

aaccgccttc gatcgtatgc tagcggggac ggtgagcagt acacgcaaaa gtcgcaatct 2160

gggagcgatt acctctcgta cctgtcagca cccatttaca attctacgca cattctggcc 2220

acgcgcaagg ggtttgcggg caatcagatc gtgagtgtcg tatccaatct gggagccaag 2280

ccagccagca aagccaccac gaaaatcacg ttgggctcag acgagactgg attccaatcc 2340

aagcagaatg tgaccgagat cttatcgtgc aagacgtatg tcacggattc cagcgggaat 2400

ctggcagtgg atctgagctc ggacgggggt ccgcgcgtgt actatcccac ggacagcctg 2460

aaggacagca ctgatatctg tggtgatcag accaaatcgg cgaccccgag tagctccgca 2520

gcctcgtccg cgagcttgac ccagtccaag ggttcagaga cctgtttgtt tggggtgccg 2580

ttggggataa gcacattggt ggtcacagtt gcgatggcca cgtcctacgt gttctagtct 2640

ccttgagcct ctacccctc tgtttgtcgc ctgtctttcc ctgtttcgca ttacccctat 2700

gattattgtt tgctcataac gctcttttga atgttttgaa actgtggctt tacttgaaac 2760

ccgataccac ccctctggct gcactgatac attactaagt atagaatgtg gcgtcttgtc 2820

tcttacgagt caccctagca tcacggtaaa ataaatgtgt tgtttccgtc tgagtctctt 2880

gaaatctttt gtagcacggg ggaaagaaaa ttgtcagggt ctccttctgg cacggatggc 2940

atagacgaag cagtatacag ggaggtaaat tcgttcgggt ttcccgactg ttacgggctt 3000

t 3001

<210> 7
<211> 1485
<212> DNA
<213> Aspergillus niger

<400> 7
atgacaatct ttctgtttct ggccattttc gtggctacag ctctggcagc cacgcctgca 60

gaatggcgct cccagtcgat atatttcctg ctcaccgatc gctttgcgcg aacggataat 120

tctaccactg cttcttgtga cttgagcgct cggcaatatt gcggtggatc ctggcagggc 180

atcatcaatc agctggacta tattcaagga atgggcttta cagcgatctg gatcacaccc 240

gtaactgcac agatccccca agatactggt tacggacagg catatcacgg atactggcag 300

caggacgctt atgccctgaa ctcccattat ggtacggcag acgatctcaa agctctggct 360

tcagctcttc actcacgggg catgtatctc atggtggacg ttgttgccaa tcacatgggc 420

cacaatggta cggggagctc tgtggactac agtgtttata ggccatttaa ttcgcaaaag 480

tactttcaca acctctgttg gatctctgat tacaataacc agacaaacgt tgaagactgc 540

tggctaggcg ataacaccgt tgccttgccg gatcttgata ctaccagtac ggaggtgaag 600

aatatgtggt atgactgggt cgagtctctc gtctctaact actccgtcga cggcctccgc 660

gtagacacag tcaagaacgt acagaagaac ttctggcccg gctacaacaa tgcttcaggc 720

gtgtactgta ttggagaagt cttcgatggg gacgcctcat acacctgtcc ttatcaggaa 780

gacttggacg gagtccttaa ttaccccatg tactatccac tcctccgggc tttcgaatcc 840

accaacggca gtatcagcga cctctataac atgatcaaca ccgtgaaatc cacctgcaga 900

gattctacgc ttctagggac cttcgtcgaa aaccacgata acccacgctt tgccaaaagc 960

tacacaagcg acatgtccct agccaaaaat gccgcaacat tcactatcct ggctgacggc 1020

attcccatca tatacgccgg tcaggaacag cactatagcg gcggtaatga cccctacaac 1080

cgcgaagcga cctggctttc aggctacaag accaccagcg agctctacac ccatatcgcc 1140

gcatcgaaca agattcgcac ccacgctata aaacaggata ccggatatct cacctacaaa 1200

aactacccca tctaccaaga cacctcgacc cttgccatgc gcaaaggcta caatggcacc 1260

caaactatca cggtcctttc taaccttggc gcctcggggt cctcatacac actctccctc 1320

ccaggaacag gctacacagc cggccaaaag attactgaaa tctatacctg cacgaatcta 1380

acagtcaact caaatggctc ggtgccagta cccatgaaga gcgggttacc gcggatcctc 1440

tatcctgcag ataagttggt taatggaagc tcattttgca gttag 1485

<210> 8

<211> 1650

<212> DNA

<213> Aspergillus niger

<400> 8

atgtttcgaa aatccgcttc cctcctgggc caacggctca tggccgtttg tctcctgtgc 60

tggtgcgttt cgctagcgac cgccgcaagc acagaagaat ggaagacgcg atccatctac 120

cagacgatga cggatcggtt tgccctcacc aacggttcga cgaccgcacc atgcaatact 180

acggtagcca actactgtgg cgggtcttgg caggggacga tcgataagct ggactacatc 240

caaggcatgg ggttcgatgc gatcatgatc tctcccgtaa ttaaaaacat tgcggggcga 300

tctaaggacg gtgaggccta ccatggatac tggcctctgg atctatacga gatcaattct 360

catttcggaa cccgggagga actgttgaag ctgagcgagg agatccacgc acgcggcatg 420

tacttgctgc tggacgtcgt catcaataat atggcttaca tgacggacgg cgaggatcct 480

gcgacgacca ttgactacaa tgttttccct cagttcaatg gatcttcata cttccacccc 540

tactgtctta ttacgaactg gaataactat acggacgcgc agtggtgtca gactggtgac 600

aattatactg cactcccaga tctgtacaca gaacacaccg cggtgcagaa catcttaatg 660

gactggagta aatcggtcat cagcaattac tccgtcgacg ggctacgaat cgacgctgcc 720

aagtccctca ctcccagctt tctgcctaca tatgcgagta ccgtgggcgg attcatgacc 780

ggcgaagtca tggattcgaa cgccaccaac gtgtgcaaat atcagacgga ttacctgcca 840

agtcttccaa actacccccc ctattactcc atgatcacgg ccttcctcaa cggcgagcct 900

gcgaccttgc tcgaggaaat cgcgacgatc aatgatcttt gccctgatac gttcgcaatg 960

gtcaacttca tcgaagatca agacgttgac cgatgggcct acatgaatga cgacatcatg 1020

ctagctaaaa ctgcactgac cttcatgatg ctctacgacg gtattccctt ggtctaccag 1080

ggtctggagc aggccattgc ctattccaac cgagcggcct tgtggttgac agatttcgac 1140

accaatgcga cgctttataa acacatcaag aaactcaatg ccatccgcaa acatgccatt 1200

aaccttgatt ccagctacat cagttcgaaa acatatccca tctatcaagg aggtagcgag 1260

ttggctttct ggaagggcaa caatggacgc caagtcatca tggttttatc tacggcgggc 1320

tcgaatggtt cagcttatac tctgacacta cccgtaagct atggggctag tgaggtggtc 1380

acggaagtgc tgaactgcgt caattatacg gtcaacactt acagtcaatt ggtcgtggac 1440

atggataagg gcgaacctcg ggtcttcttt cccgcttcga tgatgccggg gagcgggctg 1500

tgtggctaca atacttctaa tgtgacgtat tctgagctga gacttgctgc tgtgggatct 1560

tcgtcgtctg ctggtagcca ctctgtcata ccgtctgctt ttgcttcgct tttcatggcg 1620

atagttgcat ttttggcatt ccggatataa 1650

<210> 9

<211> 1668

<212> DNA

<213> Aspergillus niger

<400> 9

atggtctcaa tgtcggccct gcggcacggg cttggggtcc tctatcttgc tagctggctg 60

gggtcgtcgc tggctgccag caccgagcaa tggaagtccc ggtcaatcta tcagaccatg 120

acggatcggt tcgcgcgcac tgatggctca accacctccc cctgcaacac cacggagggt 180

ctgtactgtg gtggtacctg gcgcggcatg atcaaccatc tggattacat ccaggggatg 240

ggctttgatg ccgtcatgat ctcccctatc atcgagaacg tcgaaggtcg cgtggagtac 300

ggagaagcct atcacggcta ttggcccgtg gatctgtact ccctcaactc gcactttgga 360

acccaccagg atctactgga cctgagtgat gccctgcatg ctcgcgacat gtatctgatg 420

atggacacag tcatcaacaa catggcctac atcaccaatg gctccgaccc cgccacccac 480

atcgactact caaccctgac ccccttcaat agctcgtctt actaccatcc ctactgtaag 540

atcaccgact ggaataactt caccaatgcc cagctgtgcc aaaccggtga caacatcgtg 600

gccttgcccg atctgtatac cgagcatgcc gaggtgcaag aaaccctgag caactgggcc 660

aaggaagtca tctccaccta ttccatcgac ggcctccgca tcgatgccgc aaaacatgtg 720

aaccccggct tcctgaagaa tttcggcgat gcgcttgata tcttcatgac cggcgaagtg 780

ctgcagcaag aggtcagcac aatctgcgat tatcagaaca actacatcgg cagtctcccg 840

aattaccccg tctactacgc catgttgaag gcctttacct tgggcaacac cagcgccctg 900

gccactcagg tccagtcgat gaagaattcc tgtaatgatg tgaccgcctt gtcgtcattc 960

tccgaaaacc acgatgtcgc acgatttgcc agcatgaccc atgacatggc tctcgcgaag 1020

aatattctca cattcactct tctcttcgac ggtgttccca tgatctatca aggtcaagag 1080

cagcatctcg acggccctgg cagtccggaa aaccgagaag cgatatggct ctctgagtac 1140

aacaccgacg ccgagctcta caagttgatc ggcaagttga acgccatccg gaagcacgcc 1200

taccgactcg acaatcacta ccccgatgtc gagacgtacc ccatcttcga aggtggcagt 1260

gaactgggat tccgcaaggg aatcgaaggt cgacaggtgg tcatgcttct gtccacgcaa 1320

ggcacgaaca gcagcgctta caacctctcg atgccggtga gtttcacggg tggcacagta 1380

gtgaccgaga ttctgaactg cgtcaactac acggtcaaca cccagagcga actcgtggtt 1440

ccgatggata aaggagagcc gcgcgtgttc ttcccggcgg atctgatgcc cggcagtggg 1500

ctgtgcggac ttcctgtcgc gaatgtcacc tatgctgcct tgaggacgca gggtgcagcg 1560

gcggcggagg ctgccttgtc gctaggtatc aagactgatg cagcttctag tgctttgctt 1620

tctctggggc tgtctgtggt ggcgggtctg attgtgggga tgtggtaa 1668

<210> 10

<211> 1629

<212> DNA

<213> Aspergillus niger

<400> 10
atgttcagtt tcctcccctg cttcaaaact cgccgacaac gcaccaaatc ccaaactcaa 60
tcccaaaagc aaatagagga aaaagccaac gccatagaat ccctcccctc atggcactcc 120
cccacggaaa acaccctcct atttcaagcc ttcgaatggc acgtccccgc aacccccaac 180
accgcggaca agcgcagcca ctggcgccgg ctccagcatg ccttaccagc tatccactcc 240
ctcggtgtaa ccagcatctg gataccgcct ggatgcaagg gcatggacac aaatggaaat 300
gggtatgata tctacgattt atatgatcta ggcgaattcg accagaaagg agcagtgcga 360
acgaaatggg gcacgcgcgg agaattagag gatcttgttc gagatgcaaa tgctctgggc 420
gtaggtgtat tatgggatgc ggtgttgaat cataaggctg gggcggatag cgtggagagg 480
ttcgagggac gagatatcga agatggtaat ccccaacaaa tctccggctg gacgtccttc 540
accttccccg gccgcggcac cacatacagc ccccttcaat accactggca gcacttcagc 600
ggcgtcgact gggacgacgc ccaacaacgc aaagcaatct acaagatcct cgacccttcc 660
cgccccgaca agaactgggc ccaggacgtc ggcacagacg agaacggcaa ctatgactac 720
ctgatgttcg cggacctgga cttctcacac ccggaggtcc gggaggatgt actgcgctgg 780
gggaagtgga ttatgtctgt gttgccgttg agcggaatgc ggttagatgc ggcgaagcat 840
ttctcgacgg catttcagag ggactttatt gattgtgtgc ggcaggaggc cggggacagg 900
aaggtatctg gaggagatgg agtatcgggt ggcggcggtg gatgtgccgc tggtagagag 960
gttctcgagg ttgtcgaggg tgaagagggc ggatttgagg ggagtgttga gggggacgtt 1020
ggtggagagc aggccgggga atgcgttggt gagtttatcc cagtgatcca attcgggggt 1080
ggtgaggctg atcaaccggg acaaatgctc gagaccatcg tcgagccctc cttcaaaccc 1140
ctagcctacg ccctaatcct cctccgccaa ggaggccatc catgtgtctt ctacggcgac 1200
ctctacggca cctgcgacgg cgaccacccg ccaactcccg cctgcgaggg ccaacttccg 1260
aatctgatgc gtgcccgcaa gctgtacgcc tacggtgagc aggaggacta cttcgaccag 1320
cccaactgta ttggattcat ccgctacggc aacgccgccc acccctccgg actggcctgc 1380
gtcatgagca acggcggacc agccacgaag cgcatgtacg tgggtcggaa acatgccggc 1440
gagaagtgga cggatctgct gcagcgaggc ggtgatcacc cgtctgtcac ggtgattgat 1500
gaaatggggt acggggagtt tccggttcag agtatgaggg tgagtgtctg ggtggatagt 1560

gcggcagatg gccgagaggg tgtcggggca gaatttgacg tcgatatcta cggcattcag 1620

gcactataa

<210> 11

<211> 1695

<212> DNA

<213> Aspergillus niger

<400> 11

atgttatctt tcctcctatg gtgccatcca aagaagagga aggagagaca actatggaag 60

cagattgaag aagaagctga gcacttggac cagcttccat catgggatgc tcccgacaac 120

actttaatgc tgcaagcctt tgaatggcat gtgccagctg accaaggtca ctggcgtcgt 180

cttcaccagg ccttaccaaa cttcaaggcg attggggtag acaacatatg gatcccgccc 240

gggtgtaaag ctatgaatcc atctggtaac ggatatgata tttacgattt gtatgacttg 300

ggagaattcg agcaaaaggg gtctcgagct actaaatggg gtaccaaaga ggagctccag 360

tccttggtag ctgcagccca agatttcggt ataggtattt actgggacgc tgtcctcaat 420

cacaaggcgg gagcagatta cgcagagcgc tttcaagctg tcagggttga cccacaggag 480

cgtaatatga aaatcgcccc tgcagaggaa attgaaggct gggtgggatt caacttctct 540

gggcgtggca accactatag ttcgatgaag tacaacaaaa accacttcag cggtatcgac 600

tgggaccagt cgcgtcaaaa atgcggagtc tacaagatcc aaggacatga atgggcgaac 660

gacgtcgcca atgagaacgg aaattacgac tatctcatgt tcgccaactt ggactactcc 720

aacgcagaag tacgacgcga tgttctgaaa tgggccgagt ggctcaatgc tcaattgcct 780

ctaagcggca tgaggttgga tgcggtcaaa cattactcgg ctggttttca gaaagagctt 840

attgatcatc ttcgaactat tgctgggcca gactatttca tagtgggcga gtactggaaa 900

ggcgagacca agccgttagt tgactacctg aagcagatgg actacaagct atcattgttc 960

gattccgctc tggttgggcg gttctcaagc atttcacaga caccaggggc ggatcttcgc 1020

aacattttct ataatacatt ggtccaattg tacccagatc attctgtcgt gcaatatgcc 1080

ctttggcctc attctactaa catagatcca aagcaaccag gtcaatccct cgaagcgcca 1140

gtaacatcat tcttcaaacc cctcgcgtac gcccttatcc ttctccgtga ccaagggcaa 1200

ccatgcatat tctacggaga cctttacggc ctccaagccg atgtcaaaga tccaatgaca 1260

ccgtcttgca ggggcaagct gtccatcctc acccgagctc gaaagctcta cgcatatggc 1320

ctgcaacgag attattttga caagccgaac tgcatcggtt ttgtccgcta tggtaaccgt 1380

cggcatccct ctggtcttgc atgcgtgatg agcaatgcgg gtccgtcgag gaagcggatg 1440

tatgttggtc gacgacacgc caagcaaaca tggaccgata tcctgcagtg gtgtgatcag 1500

actgttgtca ttgatgccaa gggatatgga gagtttccgg ttagtgcgat gagcgtgagt 1560

gtatgggtga actccgaggc tgaggggaga gatagcctct cacatcattt gtatgtccct 1620

gctcacagtc ttagtacgga tgatttgctg acgagtgctg aaattagtga tgagaacata 1680

tacaaactag cttaa 1695

<210> 12

<211> 1704

<212> DNA

<213> Aspergillus niger

<400> 12

atggacgacg ggtggtggat catctcgttg ctggtggtaa ctctgggaat ccccacggtt 60

aatgcggctt ccagagacca atggatcggt cgctctatct atcaaatcgt gaccgatcgc 120

tttgctcgat cggataactc aaccaccgct gcttgtgatg cagcactagg aaactactgc 180

ggaggctctt tccagggcat catcaataag ctggactaca tccaggagct cgggtttgat 240

gcgatatgga tctctcccgc acaaagccaa atttccgccc gaacagcaga tctctcagca 300

taccatggat attggcccaa tgatctgtat tccatcaact ctcattttgg cactcccaag 360

gagctggaag ccttgtcctc tgccctgcac gatcgtggca tgtacttgat gcttgacatc 420

gtggttggtg atatggcctg ggcgggaaat cacagcaccg tcgattacag caactttaat 480

cctttcaatg atcagaagtt tttccatgat ttcaagctcc tctccagtga ccccacaaat 540

gaaacttgtg ttctggattg ctggatggga gacaccgtcg tatcacttcc cgatctgcga 600

aatgaagacc aacaagttca gaatattctt ggtacctgga tttcggggtt ggtttcgaac 660

tactcaattg acggactgcg tattgacagc gtcttaaata tcgccccgga cttcttctcc 720

aacttcacca agtcatcagg ggttttcact gtcggtgaag gtgccacggc cgatgcggct 780

gatgtttgcc ctctgcagcc aagtttaaat gggcttttga attatccatt gtactatatt 840

cttaccgacg ccttcaacac gaccaacggg aacctgagca ccattaccga gtccataagc 900

tacaccaaag gacagtgcga ggatgtcttg gctctgggga cgtttactgc aaaccaggat 960

gttccgcgct tcggttcata cacgtcggat atttcgctgg cgcgcaatat cctgaccagc 1020

agtatgctga cggacggcat ccccatcctc tattacggcg aggagcagca tttgacagga 1080

tcctacaatc ctgtcaaccg tgaggctctg tggctgacca actactcgat gcgctcgacc 1140

tccctcccca ccctcgtcca atccctgaac cgccttcgat cgtatgctag cggggacggt 1200

gagcagtaca cgcaaaagtc gcaatctggg agcgattacc tctcgtacct gtcagcaccc 1260

atttacaatt ctacgcacat tctggccacg cgcaaggggt ttgcgggcaa tcagatcgtg 1320

agtgtcgtat ccaatctggg agccaagcca gccagcaaag ccaccacgaa aatcacgttg 1380

ggctcagacg agactggatt ccaatccaag cagaatgtga ccgagatctt atcgtgcaag 1440

acgtatgtca cggattccag cgggaatctg gcagtggatc tgagctcgga cgggggtccg 1500

cgcgtgtact atcccacgga cagcctgaag gacagcactg atatctgtgg tgatcagacc 1560

aaatcggcga ccccgagtag ctccgcagcc tcgtccgcga gcttgaccca gtccaagggt 1620

tcagagacct gtttgtttgg ggtgccgttg gggataagca cattggtggt cacagttgcg 1680

atggccacgt cctacgtgtt ctag 1704

<210> 13
<211> 494
<212> PRT
<213> Aspergillus niger

<400> 13

Met Thr Ile Phe Leu Phe Leu Ala Ile Phe Val Ala Thr Ala Leu Ala
1 5 10 15

Ala Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile Tyr Phe Leu Leu Thr
20 25 30

Asp Arg Phe Ala Arg Thr Asp Asn Ser Thr Thr Ala Ser Cys Asp Leu
35 40 45

Ser Ala Arg Gln Tyr Cys Gly Gly Ser Trp Gln Gly Ile Ile Asn Gln
50 55 60

Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile Trp Ile Thr Pro
65 70 75 80

Val Thr Ala Gln Ile Pro Gln Asp Thr Gly Tyr Gly Gln Ala Tyr His
85 90 95

Gly Tyr Trp Gln Gln Asp Ala Tyr Ala Leu Asn Ser His Tyr Gly Thr
100 105 110

Ala Asp Asp Leu Lys Ala Leu Ala Ser Ala Leu His Ser Arg Gly Met
115 120 125

Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly His Asn Gly Thr
130 135 140

Gly Ser Ser Val Asp Tyr Ser Val Tyr Arg Pro Phe Asn Ser Gln Lys
145 150 155 160

Tyr Phe His Asn Leu Cys Trp Ile Ser Asp Tyr Asn Asn Gln Thr Asn
165 170 175

Val Glu Asp Cys Trp Leu Gly Asp Asn Thr Val Ala Leu Pro Asp Leu
180 185 190

Asp Thr Thr Ser Thr Glu Val Lys Asn Met Trp Tyr Asp Trp Val Glu
195 200 205

Ser Leu Val Ser Asn Tyr Ser Val Asp Gly Leu Arg Val Asp Thr Val
210 215 220

Lys Asn Val Gln Lys Asn Phe Trp Pro Gly Tyr Asn Asn Ala Ser Gly
225 230 235 240

Val Tyr Cys Ile Gly Glu Val Phe Asp Gly Asp Ala Ser Tyr Thr Cys
245 250 255

Pro Tyr Gln Glu Asp Leu Asp Gly Val Leu Asn Tyr Pro Met Tyr Tyr
260 265 270

Pro Leu Leu Arg Ala Phe Glu Ser Thr Asn Gly Ser Ile Ser Asp Leu
275 280 285

Tyr Asn Met Ile Asn Thr Val Lys Ser Thr Cys Arg Asp Ser Thr Leu
290 295 300

Leu Gly Thr Phe Val Glu Asn His Asp Asn Pro Arg Phe Ala Lys Ser
305 310 315 320

Tyr Thr Ser Asp Met Ser Leu Ala Lys Asn Ala Ala Thr Phe Thr Ile
325 330 335

Leu Ala Asp Gly Ile Pro Ile Ile Tyr Ala Gly Gln Glu Gln His Tyr
340 345 350

Ser Gly Gly Asn Asp Pro Tyr Asn Arg Glu Ala Thr Trp Leu Ser Gly
355 360 365

Tyr Lys Thr Thr Ser Glu Leu Tyr Thr His Ile Ala Ala Ser Asn Lys
370 375 380

Ile Arg Thr His Ala Ile Lys Gln Asp Thr Gly Tyr Leu Thr Tyr Lys
385 390 395 400

Asn Tyr Pro Ile Tyr Gln Asp Thr Ser Thr Leu Ala Met Arg Lys Gly
405 410 415

Tyr Asn Gly Thr Gln Thr Ile Thr Val Leu Ser Asn Leu Gly Ala Ser
420 425 430

Gly Ser Ser Tyr Thr Leu Ser Leu Pro Gly Thr Gly Tyr Thr Ala Gly
435 440 445

Gln Lys Ile Thr Glu Ile Tyr Thr Cys Thr Asn Leu Thr Val Asn Ser
450 455 460

Asn Gly Ser Val Pro Val Pro Met Lys Ser Gly Leu Pro Arg Ile Leu
465 470 475 480

Tyr Pro Ala Asp Lys Leu Val Asn Gly Ser Ser Phe Cys Ser
485 490

<210> 14

<211> 549

<212> PRT

<213> Aspergillus niger

<400> 14

Met Phe Arg Lys Ser Ala Ser Leu Leu Gly Gln Arg Leu Met Ala Val
1 5 10 15

Cys Leu Leu Cys Trp Cys Val Ser Leu Ala Thr Ala Ala Ser Thr Glu
20 25 30

Glu Trp Lys Thr Arg Ser Ile Tyr Gln Thr Met Thr Asp Arg Phe Ala
35 40 45

Leu Thr Asn Gly Ser Thr Thr Ala Pro Cys Asn Thr Thr Val Ala Asn
50 55 60

Tyr Cys Gly Gly Ser Trp Gln Gly Thr Ile Asp Lys Leu Asp Tyr Ile
65 70 75 80

Gln Gly Met Gly Phe Asp Ala Ile Met Ile Ser Pro Val Ile Lys Asn
85 90 95

Ile Ala Gly Arg Ser Lys Asp Gly Glu Ala Tyr His Gly Tyr Trp Pro
100 105 110

Leu Asp Leu Tyr Glu Ile Asn Ser His Phe Gly Thr Arg Glu Glu Leu
115 120 125

Leu Lys Leu Ser Glu Glu Ile His Ala Arg Gly Met Tyr Leu Leu Leu
130 135 140

Asp Val Val Ile Asn Asn Met Ala Tyr Met Thr Asp Gly Glu Asp Pro
145 150 155 160

Ala Thr Thr Ile Asp Tyr Asn Val Phe Pro Gln Phe Asn Gly Ser Ser
165 170 175

Tyr Phe His Pro Tyr Cys Leu Ile Thr Asn Trp Asn Asn Tyr Thr Asp
180 185 190

Ala Gln Trp Cys Gln Thr Gly Asp Asn Tyr Thr Ala Leu Pro Asp Leu
195 200 205

Tyr Thr Glu His Thr Ala Val Gln Asn Ile Leu Met Asp Trp Ser Lys

210 215 220

Ser Val Ile Ser Asn Tyr Ser Val Asp Gly Leu Arg Ile Asp Ala Ala
225 230 235 240

Lys Ser Leu Thr Pro Ser Phe Leu Pro Thr Tyr Ala Ser Thr Val Gly
245 250 255

Gly Phe Met Thr Gly Glu Val Met Asp Ser Asn Ala Thr Asn Val Cys
260 265 270

Lys Tyr Gln Thr Asp Tyr Leu Pro Ser Leu Pro Asn Tyr Pro Leu Tyr
275 280 285

Tyr Ser Met Ile Thr Ala Phe Leu Asn Gly Glu Pro Ala Thr Leu Leu
290 295 300

Glu Glu Ile Ala Thr Ile Asn Asp Leu Cys Pro Asp Thr Phe Ala Met
305 310 315 320

Val Asn Phe Ile Glu Asp Gln Asp Val Asp Arg Trp Ala Tyr Met Asn
325 330 335

Asp Asp Ile Met Leu Ala Lys Thr Ala Leu Thr Phe Met Met Leu Tyr
340 345 350

Asp Gly Ile Pro Leu Val Tyr Gln Gly Leu Glu Gln Ala Ile Ala Tyr
355 360 365

Ser Asn Arg Ala Ala Leu Trp Leu Thr Asp Phe Asp Thr Asn Ala Thr
370 375 380

Leu Tyr Lys His Ile Lys Lys Leu Asn Ala Ile Arg Lys His Ala Ile
385 390 395 400

Asn Leu Asp Ser Ser Tyr Ile Ser Ser Lys Thr Tyr Pro Ile Tyr Gln
405 410 415

Gly Gly Ser Glu Leu Ala Phe Trp Lys Gly Asn Asn Gly Arg Gln Val
420 425 430

Ile Met Val Leu Ser Thr Ala Gly Ser Asn Gly Ser Ala Tyr Thr Leu
435 440 445

Thr Leu Pro Val Ser Tyr Gly Ala Ser Glu Val Val Thr Glu Val Leu
450 455 460

Asn Cys Val Asn Tyr Thr Val Asn Thr Tyr Ser Gln Leu Val Val Asp
465 470 475 480

Met Asp Lys Gly Glu Pro Arg Val Phe Phe Pro Ala Ser Met Met Pro
485 490 495

Gly Ser Gly Leu Cys Gly Tyr Asn Thr Ser Asn Val Thr Tyr Ser Glu
500 505 510

Leu Arg Leu Ala Ala Val Gly Ser Ser Ser Ser Ala Gly Ser His Ser
515 520 525

Val Ile Pro Ser Ala Phe Ala Ser Leu Phe Met Ala Ile Val Ala Phe
530 535 540

Leu Ala Phe Arg Ile
545

<210> 15

<211> 555

<212> PRT

<213> Aspergillus niger

<400> 15

Met Val Ser Met Ser Ala Leu Arg His Gly Leu Gly Val Leu Tyr Leu
1 5 10 15

Ala Ser Trp Leu Gly Ser Ser Leu Ala Ala Ser Thr Glu Gln Trp Lys
20 25 30

Ser Arg Ser Ile Tyr Gln Thr Met Thr Asp Arg Phe Ala Arg Thr Asp
35 40 45

Gly Ser Thr Thr Ser Pro Cys Asn Thr Thr Glu Gly Leu Tyr Cys Gly
50 55 60

Gly Thr Trp Arg Gly Met Ile Asn His Leu Asp Tyr Ile Gln Gly Met
65 70 75 80

Gly Phe Asp Ala Val Met Ile Ser Pro Ile Ile Glu Asn Val Glu Gly
85 90 95

Arg Val Glu Tyr Gly Glu Ala Tyr His Gly Tyr Trp Pro Val Asp Leu
100 105 110

Tyr Ser Leu Asn Ser His Phe Gly Thr His Gln Asp Leu Leu Asp Leu
115 120 125

Ser Asp Ala Leu His Ala Arg Asp Met Tyr Leu Met Met Asp Thr Val
130 135 140

Ile Asn Asn Met Ala Tyr Ile Thr Asn Gly Ser Asp Pro Ala Thr His
145 150 155 160

Ile Asp Tyr Ser Thr Leu Thr Pro Phe Asn Ser Ser Ser Tyr Tyr His
165 170 175

Pro Tyr Cys Lys Ile Thr Asp Trp Asn Asn Phe Thr Asn Ala Gln Leu
180 185 190

Cys Gln Thr Gly Asp Asn Ile Val Ala Leu Pro Asp Leu Tyr Thr Glu
195 200 205

His Ala Glu Val Gln Glu Thr Leu Ser Asn Trp Ala Lys Glu Val Ile
210 215 220

Ser Thr Tyr Ser Ile Asp Gly Leu Arg Ile Asp Ala Ala Lys His Val
225 230 235 240

Asn Pro Gly Phe Leu Lys Asn Phe Gly Asp Ala Leu Asp Ile Phe Met
245 250 255

Thr Gly Glu Val Leu Gln Gln Glu Val Ser Thr Ile Cys Asp Tyr Gln
260 265 270

Asn Asn Tyr Ile Gly Ser Leu Pro Asn Tyr Pro Val Tyr Tyr Ala Met
275 280 285

```
Leu Lys Ala Phe Thr Leu Gly Asn Thr Ser Ala Leu Ala Thr Gln Val
    290                 295                 300

Gln Ser Met Lys Asn Ser Cys Asn Asp Val Thr Ala Leu Ser Ser Phe
305                 310                 315                 320

Ser Glu Asn His Asp Val Ala Arg Phe Ala Ser Met Thr His Asp Met
                325                 330                 335

Ala Leu Ala Lys Asn Ile Leu Thr Phe Thr Leu Leu Phe Asp Gly Val
            340                 345                 350

Pro Met Ile Tyr Gln Gly Gln Glu Gln His Leu Asp Gly Pro Gly Ser
        355                 360                 365

Pro Glu Asn Arg Glu Ala Ile Trp Leu Ser Glu Tyr Asn Thr Asp Ala
    370                 375                 380

Glu Leu Tyr Lys Leu Ile Gly Lys Leu Asn Ala Ile Arg Lys His Ala
385                 390                 395                 400

Tyr Arg Leu Asp Asn His Tyr Pro Asp Val Glu Thr Tyr Pro Ile Phe
                405                 410                 415

Glu Gly Gly Ser Glu Leu Gly Phe Arg Lys Gly Ile Glu Gly Arg Gln
            420                 425                 430

Val Val Met Leu Leu Ser Thr Gln Gly Thr Asn Ser Ser Ala Tyr Asn
        435                 440                 445

Leu Ser Met Pro Val Ser Phe Thr Gly Gly Thr Val Val Thr Glu Ile
    450                 455                 460

Leu Asn Cys Val Asn Tyr Thr Val Asn Thr Gln Ser Glu Leu Val Val
465                 470                 475                 480

Pro Met Asp Lys Gly Glu Pro Arg Val Phe Phe Pro Ala Asp Leu Met
                485                 490                 495

Pro Gly Ser Gly Leu Cys Gly Leu Pro Val Ala Asn Val Thr Tyr Ala
            500                 505                 510

Ala Leu Arg Thr Gln Gly Ala Ala Ala Ala Glu Ala Ala Leu Ser Leu
```

515 520 525

Gly Ile Lys Thr Asp Ala Ala Ser Ser Ala Leu Leu Ser Leu Gly Leu
530 535 540

Ser Val Val Ala Gly Leu Ile Val Gly Met Trp
545 550 555

<210> 16

<211> 542

<212> PRT

<213> Aspergillus niger

<400> 16

Met Phe Ser Phe Leu Pro Cys Phe Lys Thr Arg Arg Gln Arg Thr Lys
1 5 10 15

Ser Gln Thr Gln Ser Gln Lys Gln Ile Glu Glu Lys Ala Asn Ala Ile
20 25 30

Glu Ser Leu Pro Ser Trp His Ser Pro Thr Glu Asn Thr Leu Leu Phe
35 40 45

Gln Ala Phe Glu Trp His Val Pro Ala Thr Pro Asn Thr Ala Asp Lys
50 55 60

Arg Ser His Trp Arg Arg Leu Gln His Ala Leu Pro Ala Ile His Ser
65 70 75 80

Leu Gly Val Thr Ser Ile Trp Ile Pro Pro Gly Cys Lys Gly Met Asp
85 90 95

Thr Asn Gly Asn Gly Tyr Asp Ile Tyr Asp Leu Tyr Asp Leu Gly Glu
100 105 110

Phe Asp Gln Lys Gly Ala Val Arg Thr Lys Trp Gly Thr Arg Gly Glu
115 120 125

Leu Glu Asp Leu Val Arg Asp Ala Asn Ala Leu Gly Val Gly Val Leu
130 135 140

```
Trp Asp Ala Val Leu Asn His Lys Ala Gly Ala Asp Ser Val Glu Arg
145                 150                 155                 160

Phe Glu Gly Arg Asp Ile Glu Asp Gly Asn Pro Gln Gln Ile Ser Gly
                165                 170                 175

Trp Thr Ser Phe Thr Phe Pro Gly Arg Gly Thr Thr Tyr Ser Pro Leu
            180                 185                 190

Gln Tyr His Trp Gln His Phe Ser Gly Val Asp Trp Asp Asp Ala Gln
        195                 200                 205

Gln Arg Lys Ala Ile Tyr Lys Ile Leu Asp Pro Ser Arg Pro Asp Lys
    210                 215                 220

Asn Trp Ala Gln Asp Val Gly Thr Asp Glu Asn Gly Asn Tyr Asp Tyr
225                 230                 235                 240

Leu Met Phe Ala Asp Leu Asp Phe Ser His Pro Glu Val Arg Glu Asp
                245                 250                 255

Val Leu Arg Trp Gly Lys Trp Ile Met Ser Val Leu Pro Leu Ser Gly
            260                 265                 270

Met Arg Leu Asp Ala Ala Lys His Phe Ser Thr Ala Phe Gln Arg Asp
        275                 280                 285

Phe Ile Asp Cys Val Arg Gln Glu Ala Gly Asp Arg Lys Val Ser Gly
    290                 295                 300

Gly Asp Gly Val Ser Gly Gly Gly Gly Gly Cys Ala Ala Gly Arg Glu
305                 310                 315                 320

Val Leu Glu Val Val Glu Gly Glu Glu Gly Gly Phe Glu Gly Ser Val
                325                 330                 335

Glu Gly Asp Val Gly Gly Glu Gln Ala Gly Glu Cys Val Gly Glu Phe
            340                 345                 350

Ile Pro Val Ile Gln Phe Gly Gly Gly Glu Ala Asp Gln Pro Gly Gln
        355                 360                 365
```

Met Leu Glu Thr Ile Val Glu Pro Ser Phe Lys Pro Leu Ala Tyr Ala
370 375 380

Leu Ile Leu Leu Arg Gln Gly Gly His Pro Cys Val Phe Tyr Gly Asp
385 390 395 400

Leu Tyr Gly Thr Cys Asp Gly Asp His Pro Pro Thr Pro Ala Cys Glu
405 410 415

Gly Gln Leu Pro Asn Leu Met Arg Ala Arg Lys Leu Tyr Ala Tyr Gly
420 425 430

Glu Gln Glu Asp Tyr Phe Asp Gln Pro Asn Cys Ile Gly Phe Ile Arg
435 440 445

Tyr Gly Asn Ala Ala His Pro Ser Gly Leu Ala Cys Val Met Ser Asn
450 455 460

Gly Gly Pro Ala Thr Lys Arg Met Tyr Val Gly Arg Lys His Ala Gly
465 470 475 480

Glu Lys Trp Thr Asp Leu Leu Gln Arg Gly Gly Asp His Pro Ser Val
485 490 495

Thr Val Ile Asp Glu Met Gly Tyr Gly Glu Phe Pro Val Gln Ser Met
500 505 510

Arg Val Ser Val Trp Val Asp Ser Ala Ala Asp Gly Arg Glu Gly Val
515 520 525

Gly Ala Glu Phe Asp Val Asp Ile Tyr Gly Ile Gln Ala Leu
530 535 540

<210> 17

<211> 564

<212> PRT

<213> Aspergillus niger

<400> 17

```
Met Leu Ser Phe Leu Leu Trp Cys His Pro Lys Lys Arg Lys Glu Arg
1               5                  10                  15

Gln Leu Trp Lys Gln Ile Glu Glu Glu Ala Glu His Leu Asp Gln Leu
            20                  25                  30

Pro Ser Trp Asp Ala Pro Asp Asn Thr Leu Met Leu Gln Ala Phe Glu
        35                  40                  45

Trp His Val Pro Ala Asp Gln Gly His Trp Arg Arg Leu His Gln Ala
    50                  55                  60

Leu Pro Asn Phe Lys Ala Ile Gly Val Asp Asn Ile Trp Ile Pro Pro
65                  70                  75                  80

Gly Cys Lys Ala Met Asn Pro Ser Gly Asn Gly Tyr Asp Ile Tyr Asp
                85                  90                  95

Leu Tyr Asp Leu Gly Glu Phe Glu Gln Lys Gly Ser Arg Ala Thr Lys
            100                 105                 110

Trp Gly Thr Lys Glu Glu Leu Gln Ser Leu Val Ala Ala Ala Gln Asp
        115                 120                 125

Phe Gly Ile Gly Ile Tyr Trp Asp Ala Val Leu Asn His Lys Ala Gly
    130                 135                 140

Ala Asp Tyr Ala Glu Arg Phe Gln Ala Val Arg Val Asp Pro Gln Glu
145                 150                 155                 160

Arg Asn Met Lys Ile Ala Pro Ala Glu Glu Ile Glu Gly Trp Val Gly
                165                 170                 175

Phe Asn Phe Ser Gly Arg Gly Asn His Tyr Ser Ser Met Lys Tyr Asn
            180                 185                 190

Lys Asn His Phe Ser Gly Ile Asp Trp Asp Gln Ser Arg Gln Lys Cys
        195                 200                 205

Gly Val Tyr Lys Ile Gln Gly His Glu Trp Ala Asn Asp Val Ala Asn
    210                 215                 220

Glu Asn Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asn Leu Asp Tyr Ser
```

225 230 235 240

Asn Ala Glu Val Arg Arg Asp Val Leu Lys Trp Ala Glu Trp Leu Asn
245 250 255

Ala Gln Leu Pro Leu Ser Gly Met Arg Leu Asp Ala Val Lys His Tyr
260 265 270

Ser Ala Gly Phe Gln Lys Glu Leu Ile Asp His Leu Arg Thr Ile Ala
275 280 285

Gly Pro Asp Tyr Phe Ile Val Gly Glu Tyr Trp Lys Gly Glu Thr Lys
290 295 300

Pro Leu Val Asp Tyr Leu Lys Gln Met Asp Tyr Lys Leu Ser Leu Phe
305 310 315 320

Asp Ser Ala Leu Val Gly Arg Phe Ser Ser Ile Ser Gln Thr Pro Gly
325 330 335

Ala Asp Leu Arg Asn Ile Phe Tyr Asn Thr Leu Val Gln Leu Tyr Pro
340 345 350

Asp His Ser Val Val Gln Tyr Ala Leu Trp Pro His Ser Thr Asn Ile
355 360 365

Asp Pro Lys Gln Pro Gly Gln Ser Leu Glu Ala Pro Val Thr Ser Phe
370 375 380

Phe Lys Pro Leu Ala Tyr Ala Leu Ile Leu Leu Arg Asp Gln Gly Gln
385 390 395 400

Pro Cys Ile Phe Tyr Gly Asp Leu Tyr Gly Leu Gln Ala Asp Val Lys
405 410 415

Asp Pro Met Thr Pro Ser Cys Arg Gly Lys Leu Ser Ile Leu Thr Arg
420 425 430

Ala Arg Lys Leu Tyr Ala Tyr Gly Leu Gln Arg Asp Tyr Phe Asp Lys
435 440 445

Pro Asn Cys Ile Gly Phe Val Arg Tyr Gly Asn Arg Arg His Pro Ser
450 455 460

Gly Leu Ala Cys Val Met Ser Asn Ala Gly Pro Ser Arg Lys Arg Met
465 470 475 480

Tyr Val Gly Arg Arg His Ala Lys Gln Thr Trp Thr Asp Ile Leu Gln
485 490 495

Trp Cys Asp Gln Thr Val Val Ile Asp Ala Lys Gly Tyr Gly Glu Phe
500 505 510

Pro Val Ser Ala Met Ser Val Ser Val Trp Val Asn Ser Glu Ala Glu
515 520 525

Gly Arg Asp Ser Leu Ser His His Leu Tyr Val Pro Ala His Ser Leu
530 535 540

Ser Thr Asp Asp Leu Leu Thr Ser Ala Glu Ile Ser Asp Glu Asn Ile
545 550 555 560

Tyr Lys Leu Ala

<210> 18

<211> 567

<212> PRT

<213> Aspergillus niger

<400> 18

Met Asp Asp Gly Trp Trp Ile Ile Ser Leu Leu Val Val Thr Leu Gly
1 5 10 15

Ile Pro Thr Val Asn Ala Ala Ser Arg Asp Gln Trp Ile Gly Arg Ser
20 25 30

Ile Tyr Gln Ile Val Thr Asp Arg Phe Ala Arg Ser Asp Asn Ser Thr
35 40 45

Thr Ala Ala Cys Asp Ala Ala Leu Gly Asn Tyr Cys Gly Gly Ser Phe
50 55 60

```
Gln Gly Ile Ile Asn Lys Leu Asp Tyr Ile Gln Glu Leu Gly Phe Asp
65                  70                  75                  80

Ala Ile Trp Ile Ser Pro Ala Gln Ser Gln Ile Ser Ala Arg Thr Ala
                85                  90                  95

Asp Leu Ser Ala Tyr His Gly Tyr Trp Pro Asn Asp Leu Tyr Ser Ile
            100                 105                 110

Asn Ser His Phe Gly Thr Pro Lys Glu Leu Glu Ala Leu Ser Ser Ala
        115                 120                 125

Leu His Asp Arg Gly Met Tyr Leu Met Leu Asp Ile Val Val Gly Asp
    130                 135                 140

Met Ala Trp Ala Gly Asn His Ser Thr Val Asp Tyr Ser Asn Phe Asn
145                 150                 155                 160

Pro Phe Asn Asp Gln Lys Phe Phe His Asp Phe Lys Leu Leu Ser Ser
                165                 170                 175

Asp Pro Thr Asn Glu Thr Cys Val Leu Asp Cys Trp Met Gly Asp Thr
            180                 185                 190

Val Val Ser Leu Pro Asp Leu Arg Asn Glu Asp Gln Gln Val Gln Asn
        195                 200                 205

Ile Leu Gly Thr Trp Ile Ser Gly Leu Val Ser Asn Tyr Ser Ile Asp
    210                 215                 220

Gly Leu Arg Ile Asp Ser Val Leu Asn Ile Ala Pro Asp Phe Phe Ser
225                 230                 235                 240

Asn Phe Thr Lys Ser Ser Gly Val Phe Thr Val Gly Glu Gly Ala Thr
                245                 250                 255

Ala Asp Ala Ala Asp Val Cys Pro Leu Gln Pro Ser Leu Asn Gly Leu
            260                 265                 270

Leu Asn Tyr Pro Leu Tyr Tyr Ile Leu Thr Asp Ala Phe Asn Thr Thr
        275                 280                 285

Asn Gly Asn Leu Ser Thr Ile Thr Glu Ser Ile Ser Tyr Thr Lys Gly
```

```
    290                 295                 300

Gln Cys Glu Asp Val Leu Ala Leu Gly Thr Phe Thr Ala Asn Gln Asp
305                 310                 315                 320

Val Pro Arg Phe Gly Ser Tyr Thr Ser Asp Ile Ser Leu Ala Arg Asn
                325                 330                 335

Ile Leu Thr Ser Ser Met Leu Thr Asp Gly Ile Pro Ile Leu Tyr Tyr
            340                 345                 350

Gly Glu Glu Gln His Leu Thr Gly Ser Tyr Asn Pro Val Asn Arg Glu
        355                 360                 365

Ala Leu Trp Leu Thr Asn Tyr Ser Met Arg Ser Thr Ser Leu Pro Thr
    370                 375                 380

Leu Val Gln Ser Leu Asn Arg Leu Arg Ser Tyr Ala Ser Gly Asp Gly
385                 390                 395                 400

Glu Gln Tyr Thr Gln Lys Ser Gln Ser Gly Ser Asp Tyr Leu Ser Tyr
                405                 410                 415

Leu Ser Ala Pro Ile Tyr Asn Ser Thr His Ile Leu Ala Thr Arg Lys
            420                 425                 430

Gly Phe Ala Gly Asn Gln Ile Val Ser Val Val Ser Asn Leu Gly Ala
        435                 440                 445

Lys Pro Ala Ser Lys Ala Thr Thr Lys Ile Thr Leu Gly Ser Asp Glu
    450                 455                 460

Thr Gly Phe Gln Ser Lys Gln Asn Val Thr Glu Ile Leu Ser Cys Lys
465                 470                 475                 480

Thr Tyr Val Thr Asp Ser Ser Gly Asn Leu Ala Val Asp Leu Ser Ser
                485                 490                 495

Asp Gly Gly Pro Arg Val Tyr Tyr Pro Thr Asp Ser Leu Lys Asp Ser
            500                 505                 510

Thr Asp Ile Cys Gly Asp Gln Thr Lys Ser Ala Thr Pro Ser Ser Ser
        515                 520                 525
```

```
Ala Ala Ser Ser Ala Ser Leu Thr Gln Ser Lys Gly Ser Glu Thr Cys
    530                 535                 540

Leu Phe Gly Val Pro Leu Gly Ile Ser Thr Leu Val Val Thr Val Ala
545                 550                 555                 560

Met Ala Thr Ser Tyr Val Phe
```

CLAIMS

1) An isolated polynucleotide hybridisable to a polynucleotide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12

2) An isolated polynucleotide according to claim 1 hybridisable under high stringency conditions to a polynucleotide selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12.

3) An isolated polynucleotide according to claims 1 or 2 obtainable from a filamentous fungus.

4) An isolated polynucleotide according to claim 3 obtainable from A. niger.

5) An isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

6) An isolated polynucleotide encoding at least one functional domain of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

7) An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12 or functional equivalents thereof

8) An isolated polynucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO:

5 or SEQ ID NO: 6 or from the group consisting of SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 or SEQ ID NO: 12.

9) A vector comprising a polynucleotide sequence according to claims 1 to 8.

10) A vector according to claim 9 wherein said polynucleotide sequence according to claims 1 to 8 is operatively linked with regulatory sequences suitable for expression of said polynucleotide sequence in a suitable host cell.

11) A vector according to claim 10 wherein said suitable host cell is a filamentous fungus

12) A method for manufacturing a polynucleotide according to claims 1 – 8 or a vector according to claims 9 to 11 comprising the steps of culturing a host cell transformed with said polynucleotide or said vector and isolating said polynucleotide or said vector from said host cell.

13) An isolated polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18 or functional equivalents thereof.

14) An isolated polypeptide according to claim 15 obtainable from Aspergillus niger

15) An isolated polypeptide obtainable by expressing a polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11 in an appropriate host cell, e.g. Aspergillus niger.

16) Recombinant amylase comprising a functional domain of any of the polypeptides according to claims 13 to 15.

17) A method for manufacturing a polypeptide according to claims 13 to 16 comprising the steps of transforming a suitable host cell with an isolated polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11, culturing said cell under conditions allowing expression of said polynucleotide and optionally purifying the encoded polypeptide from said cell or culture medium.

18) A recombinant host cell comprising a polynucleotide according to claims 1 to 8 or a vector according to claims 9 to 11.

19) A recombinant host cell expressing a polypeptide according to claims 13 to 16.

20) Purified antibodies reactive with a polypeptide according to claims 13 to 16.

21) Fusion protein comprising a polypeptide sequence according to claims 13 to 16.

22) Use of an isolated polynucleotide according to claims 1-8, or a vector according to claims 9-11 or an isolated polypeptide according to claims 13-15 or a recombinant protein according to claims 16 or 21 in a baking process.